Cellular and biochemical characterisation of PrimPol, a novel eukaryotic primase-polymerase involved in DNA damage tolerance

A thesis submitted to the University of Sussex for the degree of Doctor of Philosophy

By

Sean G. Rudd

January 2013

I hereby declare that this thesis has not been and will not be, submitted in whole or in part to another University for the award of any other degree.

Sean G. Rudd

University of Sussex

Sean G. Rudd

Doctor of Philosophy Biochemistry

Cellular and biochemical characterisation of PrimPol, a novel eukaryotic primase-polymerase involved in DNA damage tolerance

Summary

Genome stability is of upmost importance to life. DNA polymerases are essential for the duplication and maintenance of the genome but they cannot themselves begin synthesis of a DNA chain, and require the activity of specialised RNA polymerases called primases. In eukaryotic cells distinct enzymes catalyse these two essential processes. This thesis contains the characterisation of coiled-coil domain containing protein (CCDC)111, a previously uncharacterised protein conserved in a broad range of unicellular and multicellular eukaryotes including humans. CCDC111 is a member of the archaeao-eukaroytic primase (AEP) superfamily and uniquely for a eukaryotic enzyme possesses both primase and polymerase activities, and was thus renamed PrimPol. The work in this thesis implicates PrimPol in the process of DNA damage tolerance, a universal mechanism by which cells complete genome duplication in spite of potentially lethal DNA damage. The first results chapters detail the essential role of a PrimPol homologue (TbPrimPol2) in the important protozoan pathogen *Trypanosoma brucei*. A combination of molecular, cell biology, and biochemical analyses indicate a role for TbPrimPol2 in the post-replication tolerance of endogenously occurring DNA damage using its trans-lesion DNA synthesis activity. The remaining results chapters characterise PrimPol in human cultured cells, and demonstrate that this enzyme is present in both the nucleus and mitochondria. In the nucleus PrimPol functions in the cellular tolerance of ultraviolet (UV)-induced DNA damage, and is required to protect xeroderma pigmentosum variant (XP-V) cells, deficient in the UV lesion bypass polymerase Pol η, from the cytotoxic affects of UV radiation. Together, this thesis establishes the involvement of PrimPol in DNA damage tolerance from one of the earliest diverging eukaryotic organisms to man.

Acknowledgements

I would first like to thank my supervisor Aidan Doherty, for offering me the opportunity to do a DPhil in his laboratory, particularly on this exciting project, and for giving me an excuse to live at the seaside these past four years. Credit is due to all the members of the Doherty lab, both past and present, for their continued tuition and support. Thanks to Julie and Andy for showing me how to function in a lab, Pierre for always having a well thought answer to my questions, Stan for his expertise, and Nigel, Laura, Helen, Ed, and Tamsyn; without them I would not have survived. I would like to thank the good people of the GDSC, who are always willing to share their time and expertise and make this a wonderful place to work; and also my recently graduated friends, John, Helen, and Ross, for plenty of advice on how to complete a PhD. I would like to thank Lucy Glover and David Horn at the London School of Hygiene and Tropical Medicine for teaching me about trypanosomes and a lesson in efficiency; and to the various people who have taught me along the way, particularly Stefan Roberts formerly at the University of Manchester, for giving me my first taste of working in a lab and inspiring me to continue. I'd like to thank my friends and family for keeping me sane during these past years, and particularly Julie Bianchi for her continued support, which without, this thesis would not have been possible; and also Colette, for putting a roof over my head these last few months.

The Medical Research Council has supported the work throughout my DPhil, and I'd also like to thank Tony Carr and Aidan for keeping me afloat when my money was running out.

Table of Contents

Abbreviations i

List of figures v

List of tables ix

Chapter 1

Introduction 1

1. Introduction 2

1.1. DNA Polymerases 2

1.1.1. Overview of eukaryotic DNA polymerases 2

1.1.2. Mechanism and structure of a DNA polymerase 3

1.1.3. Processivity and fidelity – important features of a DNA polymerase 4

1.2. Genome Duplication 6

1.2.1. Replication origins – where to build a replication fork 6

1.2.2. Initiation – biogenesis of a replication fork 7

1.2.3. Elongation – progression of the replication fork 8

1.3. Overview of DNA damage and repair 9

1.4. DNA damage tolerance 11

1.4.2. Discovery of DNA damage tolerance 11

1.4.3. Y-family DNA polymerases and Pol ζ 12

1.4.3.1. Features of Y-family DNA polymerases 12

1.4.3.2. DNA polymerase η 13

1.4.3.3. DNA polymerase ι 15

1.4.3.4. DNA polymerase κ 16

1.4.3.5. REV1 17

1.4.3.6. DNA polymerase ζ 18

1.4.4. The PCNA switchboard and the regulation of TLS 19

1.4.5. Fork-associated or post-replicative DNA damage tolerance? 21

1.5. The African Trypanosome - *Trypanosoma brucei* 22

1.5.1. African sleeping sickness 23

1.5.2. Organisation and replication of the nuclear genome 24

1.5.3. DNA repair and damage tolerance 26

1.6. DNA Primases **28**
1.6.1. Archaeal-eukaryotic primase (AEP) superfamily 28
1.6.2. Mechanism and structure of a DNA primase 29
1.6.3. Eukaryotic AEPs – priming DNA replication 31
1.6.4. Archaeal AEPs – versatile primase-polymerases? 32
1.6.5. Bacterial AEPs – NHEJ polymerases 33
1.7. PrimPol, a novel eukaryotic primase-polymerase **34**

Chapter 2

Materials and Methods **37**

2.1. Preparation of plasmid DNA **38**
2.1.1. Preparation of competent *E. coli* DH5α 38
2.1.2. Transformation of competent DH5α 38
2.1.3. Plasmid DNA amplification and purification 38
2.1.4. Agarose gel electrophoresis of DNA 39
2.2. Molecular cloning **39**
2.2.1. Polymerase chain reaction 39
2.2.2. Mutagenesis and inverse deletion PCR 40
2.2.3. Restriction digest 40
2.2.4. Ligation 41
2.2.5. Sequencing 41
2.3. Protein electrophoresis and Western blot analysis **41**
2.3.1. Sodium dodecyl sulphate polyacrylamide gel electrophoresis 41
2.3.2. Coomassie blue staining 42
2.3.3. Western blot analysis 42
2.3.4. Far-Western blot analysis of slot-blotted recombinant proteins 43
2.4. *T. brucei* methods **43**
2.4.1. Strains and maintenance 43
2.4.2. Stable transformation 44
2.4.3. Preparation of genomic DNA 44
2.4.4. Preparation of cell lysates 44
2.4.5. Immunofluorescence microscopy 44
2.4.6. TUNEL staining 45
2.4.7. Flow cytometry 45
2.5. Preparation of recombinant proteins in *E. coli* **46**
2.5.1. Preparation of chemically competent B834s 46

2.5.2. Transformation of competent B843s 46
2.5.3. Recombinant protein expression trials 46
2.5.4. Protein expression and preparation of *E. coli* lysate 47
2.5.5. Immobilised metal affinity chromatography 47
2.5.6. Heparin affinity chromatography 48
2.5.7. Storage of recombinant proteins 48
2.6. Primer extension assay 48
2.6.1. Primer-template substrates 48
2.6.2. Annealing of primer-template substrates 49
2.6.3. Primer extension assay 49
2.6.4. DNA-polyacrylamide gel electrophoresis 49
2.7. Human cell culture 50
2.7.1. Cell lines 50
2.7.2. Cell maintenance 50
2.7.3. Cell storage 51
2.8. Cultured human cell methods 51
2.8.1. Preparation of cell lysates 51
2.8.2. Stable transformation on Flp-In T-REx-293 cells 51
2.8.3. RNA interference 52
2.8.4. DNA damaging agents and drug treatments 52
2.8.5. UV-C clonogenic survival assay 53
2.8.6 Immunofluorescence microscopy 53
2.8.7. Flow cytometry 54
2.8.8. Sub-cellular fractionation 54
2.8.8.1. Mitochondrial isolation 54
2.8.8.2. Protease protection assay 55
2.8.8.3. Soluble and insoluble (chromatin) fractionation and DNase treatment 55
2.8.9. Purification of *Strep*-tagged PrimPol from human cultured cells 56
2.8.9.1 Small scale affinity purification 56
2.8.9.2. Affinity purification for mass-spectrometry analysis 56
2.8.9.3. Treatment of purified PrimPol with lambda phosphatase 57
2.9. Yeast two-hybrid methods 57
2.9.1. Yeast culture 57
2.9.2. Yeast transformation 57
2.9.3. Detection of interaction 58
2.10. Bioinformatics 58

Chapter 3

The African trypanosome contains two PrimPol-like proteins and one is essential for completion of DNA replication 60

3.1. Introduction 61

3.2. Trypanosomatid genomes encode two PrimPol-like proteins 62

3.2.1. PrimPol1 and PrimPol2 in the African trypanosome *T. brucei* 62

3.3. Generation of strains to study the cellular roles of TbPrimPol1 and 2 63

3.3.1. TbPrimPol1 and 2 inducible RNAi strains 63

3.3.2. TbPrimPol1 and 2 *in situ* tagged strains 64

3.3.3. Confirmation of TbPrimPol1 and 2 *in situ* tagged and RNAi strains 64

3.4. TbPrimPol2 is an essential protein in bloodstream form *T. brucei* 65

3.5. TbPrimPol2 is a nuclear protein 65

3.6. TbPrimPol2 is expressed in G2/M phase cells 65

3.7. TbPri, the putative replicative DNA primase in *T. brucei* 66

3.7.1. Generation of TbPriL *in situ* tagging and inducible RNAi strains 66

3.7.2 Confirmation of TbPriL *in situ* tagged and RNAi strains 66

3.7.3. TbPriL is an essential, nuclear protein in bloodstream form *T. brucei* 67

3.8. TbPrimPol2's cellular role is distinct from canonical DNA primases 67

3.8.1. Depletion of TbPrimPol2 arrests cells with 1 nucleus and 2 kinetoplasts 67

3.8.2. Depletion of TbPrimPol2 arrests cells with 4n DNA content 68

3.9. TbPrimPol2 depletion results in the accumulation of DNA damage 68

3.9.1 TbPrimPol2 depletion causes a small increase in nuclear TUNEL staining 69

3.9.2 Depletion of TbPrimPol2 triggers assembly of Rad51 foci 69

3.10. Over-expression of TbPrimPol2 does not perturb cell proliferation 70

3.11. Attempt to delete the un-tagged TbPrimPol2 allele in the *in situ* tagged strain 70

3.12. Discussion 71

Chapter 4

PrimPol1 and 2 in the African trypanosome are trans-lesion synthesis DNA polymerases 74

4.1. Introduction 75

4.2. Cloning the *T. brucei* PrimPol1 and PrimPol2 genes 75

4.3. Expression of recombinant TbPrimPol1 and 2 in *E. coli* 75

4.4. Purification of recombinant TbPrimPol1 and 2 76

4.5. Production of catalytically inactive TbPrimPol1 and 2 mutants 77
4.6. TbPrimPol1 and 2 are DNA-dependant DNA polymerases 77
4.7. TbPrimPol1 and 2 are trans-lesion synthesis DNA polymerases 78
4.7.1 Extension opposite a templated T-T CPD 78
4.7.2. Error-prone bypass of a templated T-T 6-4 photoproduct 79
4.7.3. Incorporate opposite a 3MeA analogue 80
4.7.4. Bypass of a templated 8-oxo-guanine 80
4.7.5. Stall prior to a templated abasic site 81
4.8. Discussion 83

Chapter 5

Localisation of PrimPol in human cells 87
5.1. Introduction 88
5.2. Immunofluorescent detection of PrimPol 88
5.2.1. Generation of a stable, inducible PrimPol expression cell line 89
5.2.2. Immunofluorescent detection of stably expressed recombinant PrimPol 90
5.2.3. Immunofluorescent detection of endogenous PrimPol 90
5.3. PrimPol is a mitochondrial protein 91
5.3.1. PrimPol is present in a mitochondrial preparation 91
5.3.2. PrimPol is retained in protease-treated mitochondria 91
5.4. PrimPol is a nuclear protein, associating with chromatin 92
5.5. Discussion 94
5.5.1. PrimPol – a novel mitochondrial primase-polymerase in human cells 94
5.5.2. PrimPol – a novel primase-polymerase in the nuclei of human cells 96

Chapter 6

PrimPol, a new player in DNA damage tolerance in human cells 97
6.1. Introduction 98
6.2. PrimPol re-localises in the nucleus following UV-C irradiation 98
6.2.1. PrimPol assembles into detergent-resistant foci in UV-C irradiated cells 98
6.2.2. PrimPol foci assemble in a UV-C dose-dependant manner 99
6.2.3. PrimPol foci do not assemble following exposure to ionising radiation 99
6.3. Does PrimPol re-localise to stalled replication forks? 100
6.4. PrimPol associates tightly with chromatin in UV-C irradiated cells 101
6.4.1. Recombinant PrimPol associates with chromatin in UV-C irradiated cells 101
6.4.2. Endogenous PrimPol becomes Triton insoluble in UV-C irradiated cells 102

6.5. RNAi depletion of PrimPol in normal and XP-V patient cells 102
6.5.1. PrimPol is required for the tolerance of UV-C induced DNA damage 102
6.5.2. PrimPol protects XP-V, but not normal cells, from UV-C cytotoxicity 104
6.6. PrimPol associates with chromatin at discrete foci following nucleotide deprivation 104
6.7. Discussion 106
6.9.1. UV-induced foci formation and chromatin association of PrimPol 106
6.9.2. PrimPol mediated DNA damage tolerance in normal and XP-V cells 107
6.9.3. Non UV-induced DNA damage tolerance roles of PrimPol 108
6.9.4. PrimPol – a TLS polymerase or primase, or perhaps both? 109

Chapter 7

Towards the identification of human PrimPol protein partners 111
7.1. Introduction 112
7.2. Identification of *in vivo* protein partners using the *Strep*-tag system 112
7.2.1. Over-expressed PrimPol was initially degraded 112
7.2.2 Affinity purification of *Strep*-tagged PrimPol from whole cell lysate 113
7.2.3. Mass-spectrometry analysis of *Strep*-tagged PrimPol purifications 114
7.2.4 Validating potential protein partners - RPA and mtSSB co-purify with PrimPol from human cells 114
7.2.5. PrimPol is phosphorylated *in vivo* on serine 499 and 501 115
7.3. Identification of potential protein partners *in vitro* using recombinant candidate proteins 115
7.3.1 Purification of recombinant PolG2 produced in *E. coli* 116
7.3.2. PrimPol interacts with the mtDNA replicase Pol γ *in vitro* 116
7.4. Identification of protein partners using the Yeast-Two Hybrid assay 116
7.4.1. PrimPol carboxyl-terminal containing the UL52 zinc finger auto-activates 117
7.5. Discussion 119

Chapter 8

Conclusion 121

References 125

Appendix 158

Abbreviations

3dMeA	3-deaza 3-methyladenine
3MeA	3-methyladenine
6-4 PP	Pyrimidine (6-4) pyrimidone photoproduct
8-oxo-G	8-oxo-2'-deoxyguanosine
A	Adenine
AEP	Archaeo-eukaryotic primase
APS	Ammonium persulphate
ATM	Ataxia telangiectasia mutated
ATR	Ataxia telangiectasia and Rad3 related
BER	Base excision repair
BLAST	Basic Local Alignment Search Tool
bp	base pair
BRCA2	Breast cancer type 2 susceptibility protein
BRCT	Breast cancer type 1 susceptibility protein (BRCA1) carboxyl-terminus
BSA	Bovine serum albumin
C	Cytosine
CBMSO	Centro de Biología Molecular Severo Ochoa
CCDC111	Coiled-coil domain containing 111
CDC	Cell division cycle
cDNA	Complementary DNA
Chk1	Checkpoint kinase 1
Chk2	Checkpoint kinase 2
CPD	Cyclobutane pyrimidine dimer
CRUK	Cancer Research UK
CSK	Cytoskeletal
C-terminus	Carboxyl-terminus
DAPI	4',6-diamidino-2-phenylindole
DDT	Dichlorodiphenyltrichloroethane
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DUB	De-ubiquitination enzyme
dCMP	Deoxycytidine monophosphate
dNTP	Deoxynucleotide 5' triphosphate
dRP	5'-deoxyribose phosphate
EDTA	Ethylenediaminetetra-acetic acid
EGTA	Ethyleneglycoltetra-acetic acid

FBS	Foetal bovine serum
FEN1	Flap structure-specific endonuclease 1
Flp-In T-REx	Flp Recombinase-mediated Integration Tetracycline-Regulated Expression
FRT	Flp Recombinase Target
G	Guanine
G1	Growth phase 1
G2	Growth phase 2
GDSC	Genome Damage and Stability Centre
HA	Haemagglutinin (influenza)
HEK-293	Human embryonic kidney-293
HEPES	4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid
HEX	Hexachlorofluorescein
HR	Homologous recombination
HRP	Horseradish peroxidase
Hs	*Homo sapien*
HU	Hydroxyurea
IMAC	Immobilised metal affinity chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IR	Ionising radiation
kb	Kilobase pair
kDa	Kilo Dalton
kDNA	Kinetoplast DNA
LigD	Ligase D
M	Mitosis
Mb	Megabase pair
MCM	Minichromosome maintenance
MMR	Mismatch repair
MMS	Methyl methanesulfonate
MRC	Medical Research Council
mRNA	messenger ribonucleic acid
mtDNA	Mitochondrial DNA
NER	Nucleotide excision repair
NIH	National Institutes of Health
NHEJ	Nonhomologous end-joining
NTA	Nitrilotriacetic acid
N-terminus	Amino-terminus
NTP	Nucleotide triphosphate
ORF	Open reading frame
ORC	Origin recognition complex

PAGE	Polyacrylamide gel electrophoersis
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PI	Propidium iodide
PIPES	Piperazine-N,N′-bis(2-ethanesulfonic acid)
PMSF	Phenylmethylsulfonylfluoride
PolDom	Polymerase domain of Ligase D
Pol	DNA-dependant DNA polymerase
PPi	Pyrophosphate
Pre-RC	Pre-replication complex
Prim1	DNA primase polypeptide 1
Prim2	DNA primase polypeptide 2
PriL	DNA primase large subunit
PrimPol	DNA-dependant primase-polymerase
PriS	DNA primase small subunit
PVDF	Polyvinyladine fluoride
RAD	Radiation (gene)
REV	Reversionless (gene)
RFC	Replication factor C
ROS	Reactive oxygen species
RPA	Replication protein A
RNA	Ribonucleic acid
RNAi	Ribonucleic acid interference
rNTP	Ribonucleotide triphosphate
RPE	Retinal pigment epithelial
S-phase	Synthesis-phase
SDM	Site-directed mutagenesis
SDS	Sodium dodecyl sulfate
SV40	Simian virus 40
T	Thymine
Tb	*Trypanosoma brucei*
TdT	Terminal deoxynucleotidyl transferase
TEMED	N,N,N′,N′-Tetramethylethylenediamine
Tgo	*Thermococcus gorgonarius*
TLS	Trans-lesion synthesis
Tm	Melting temperature
TUNEL	Terminal deoxynucleotidyl transferase dUTP nick end labelling
U2OS	U2-osteosarcoma
UBM	Ubiquitin-binding motif

UBZ	Ubiquitin binding zinc finger
USP-1	Ubiquitin specific protease 1
UTP	Uridine-5'-triphosphate
UTR	Un-translated region
UV	Ultraviolet
VSG	Variant surface glycoprotein
XP	Xeroderma pigmentosum
XP-A	Xeroderma pigmentosum complementation group A
XP-V	Xeroderma pigmentosum variant
XRCC1	X-ray repair cross-complementing protein 1

List of figures

Figure 1.1	The human genome encodes 15 DNA-dependant DNA polymerases	**2**
Figure 1.2	The enzymatic synthesis of DNA	**3**
Figure 1.3	Structure of a DNA polymerase	**4**
Figure 1.4	The two-metal ion mechanism of DNA polymerases	**4**
Figure 1.5	The mechanism of chromosomal DNA replication	**6**
Figure 1.6	The licensing and firing of replication origins	**6**
Figure 1.7	The initiation and elongation of DNA replication	**7**
Figure 1.8	Overview of the causes and repair mechanisms of DNA damage in human cells	**9**
Figure 1.9	Exposure of DNA to UV radiation can result in the formation of photoproducts	**9**
Figure 1.10	Overview of DNA damage tolerance mechanisms in eukaryotic cells	**11**
Figure 1.11	One and two polymerase mechanisms of trans-lesion synthesis	**12**
Figure 1.12	Regulation of DNA damage tolerance by the DNA sliding clamp PCNA	**19**
Figure 1.13	Structure of an archaeo-eukaryotic primase	**29**
Figure 1.14	The promiscuous activities of archaeal-eukaryotic primases	**32**
Figure 1.15	Domain organisation of various archaeo-eukaryotic primases	**35**
Figure 1.16	Conserved domain of the PrimPol (CCDC111) famliy	**35**
Figure 3.1	Phylogenetic analysis of PrimPol family members	**62**
Figure 3.2	Conserved motifs and domains of human PrimPol and *T. brucei* homologues	**62**
Figure 3.3	Sequence alignments of the conserved AEP and UL52 motifs from a broad range of PrimPol homologues	**62**
Figure 3.4	Alignment of the trypanosomatid PrimPol1 family	**63**
Figure 3.5	Alignment of the trypanosomatid PrimPol2 family	**63**
Figure 3.6	Generation of a tetracycline-inducible RNAi strain in *T. brucei*	**63**
Figure 3.7	Generation of an *in situ* native allele tagging strain in *T. brucei*	**64**
Figure 3.8	Validation of TbPrimPol1 and 2 *in situ* tagged and RNAi strains	**64**
Figure 3.9	TbPrimPol2 is essential whilst TbPrimPol1 is dispensable in bloodstream form *T. brucei*	**65**
Figure 3.10	TbPrimPol2 is a nuclear protein	**65**
Figure 3.11	TbPrimPol2 is predominantly detected in 1n2k (G2/M) cells	**65**
Figure 3.12	Validation of TbPriL in situ tagged and RNAi strain	**67**
Figure 3.13	TbPriL is essential in bloodstream form *T. brucei*	**67**
Figure 3.14	TbPriL is a nuclear protein	**67**

Figure 3.15	TbPrimPol2 and TbPriL are required for cell cycle progression	**67**
Figure 3.16	Abnormal cell morphology and distorted nuclei observed following TbPriL and TbPrimPol2 knockdown	**68**
Figure 3.17	TbPrimPol2 knockdown arrests cells with 4n DNA content, whilst TbPriL knockdown arrests cells during S-phase	**68**
Figure 3.18	TbPrimPol2 RNAi results in a small increase in nuclear TUNEL staining	**69**
Figure 3.19	TbPrimPol2 and TbPriL depletion triggers assembly of Rad51 foci	**69**
Figure 3.20	TbPrimPol2 over-expression has no effect on cell proliferation	**70**
Figure 3.21	An attempt to delete the un-tagged TbPrimPol2 allele in the *in situ* tagged strain	**70**
Figure 3.22	Depletion of TbPrimPol2 triggers assembly of γH2A foci	**72**
Figure 3.23	TbPrimPol2 accumulates into DNA repair foci	**72**
Figure 4.1	Cloning of TbPrimPol1 and 2 and expression in *E. coli*	**75**
Figure 4.2	Optimising expression of recombinant TbPrimPol1 and 2 in *E. coli*	**76**
Figure 4.3	Purification of recombinant TbPrimPol1 and 2 from *E. coli*	**76**
Figure 4.4	Generation of catalytically inactive mutants of TbPrimPol1 and 2	**77**
Figure 4.5	Purification of TbPrimPol1 and 2 catalytic mutants	**77**
Figure 4.6	TbPrimPol1 and 2 are DNA-dependant DNA polymerases	**77**
Figure 4.7	TbPrimPol1 and 2 cannot read-through a T-T CPD but can extend from bases opposite	**78**
Figure 4.8	TbPrimPol1 and 2 can read-through a T-T 6-4 photoproduct	**79**
Figure 4.9	TbPrimPol1 and 2 bypass of a 6-4 photoproduct is error-prone	**80**
Figure 4.10	TbPrimPol1 and 2 can incorporate opposite a 3MeA analogue	**80**
Figure 4.11	TbPrimPol1 and 2 can incorporate the correct base opposite a 3MeA analogue	**80**
Figure 4.12	TbPrimPol1 and 2 can read-through a templated 8-oxo-Guanine	**81**
Figure 4.13	TbPrimPol1 and 2 can bypass a template 8-oxo-Guanine in an error-prone and error-free manner	**81**
Figure 4.14	TbPrimPol1 and 2 cannot read-through or extend from bases opposite an abasic site	**81**
Figure 4.15	Model of the essential role of TbPrimPol2 during cell proliferation	**84**
Figure 5.1	Generation of a stable cell line with inducible expression of recombinant PrimPol	**89**
Figure 5.2	Stable, inducible over-expression of recombinant PrimPol in cultured human cells	**89**
Figure 5.3	Immunofluorescent detection of over-expressed HA-tagged PrimPol in cultured human cells	**90**

Figure 5.4	Immunofluorescent detection of PrimPol in various cultured human cell lines	**90**
Figure 5.5	Immunostaining with anti-PrimPol is specific to the PrimPol protein	**91**
Figure 5.6	PrimPol is present in a crude mitochondrial preparation	**91**
Figure 5.7	PrimPol is retained in protease treated mitochondria	**91**
Figure 5.8	PrimPol becomes insoluble following prolonged DNA replication arrest	**92**
Figure 6.1	PrimPol re-localises into sub-nuclear detergent-resistant foci following UV-C irradiation	**98**
Figure 6.2	The proportion of PrimPol focal cells was dependant upon the dose of UV-C irradiation	**99**
Figure 6.3	PrimPol did not assemble into foci following exposure to ionising radiation	**99**
Figure 6.4	PrimPol focal cells following UV-C irradiation are also focal for RPA, and a portion of these foci co-localise	**100**
Figure 6.5	UV-C induced PrimPol foci occur in S-phase cells	**100**
Figure 6.6	Over-expressed PrimPol becomes Triton insoluble following UV-C irradiation	**101**
Figure 6.7	Over-expressed PrimPol associates with chromatin following UV-C irradiation	**101**
Figure 6.8	Endogenous PrimPol becomes Triton-insoluble following UV-C irradiation	**102**
Figure 6.9	PrimPol depletion causes a proliferative defect, which is more pronounced in XP-V cells	**102**
Figure 6.10	Depletion of PrimPol results in increased chromatin-bound RPA in UV-C irradiated normal, but not XP-V, patient cells	**103**
Figure 6.11	PrimPol depletion results in over-activation of the intra-S-checkpoint in UV-C irradiated cells	**103**
Figure 6.12	PrimPol protects XP-V cells from UV cytotoxicity	**104**
Figure 6.13	PrimPol re-localises following HU treatment, assembling into foci and associating with chromatin	**104**
Figure 6.14	Possible models of PrimPol-dependant DNA damage tolerance	**106**
Figure 7.1	Schematic overview of Strep-tag affinity purification method	**112**
Figure 7.2	Stably over-expressed PrimPol was initially degraded in human cells, however only full-length was imported into mitochondria	**113**
Figure 7.3	Affinity purification of *Strep*-tagged PrimPol from whole cell lysate	**113**
Figure 7.4	Summary of proteins identified in mass-spectrometry analysis	**114**
Figure 7.5	Validating mass-spectrometry hits - PrimPol co-purifies with mtSSB and RPA	**114**

Figure 7.6	PrimPol is phosphorylated *in vivo*	**115**
Figure 7.7	Purification of the human mtDNA replicase accessory subunit PolG2 from *E. coli*	**116**
Figure 7.8	Evidence for PrimPol interaction with the mtDNA replicase Pol γ	**116**
Figure 7.9	Schematic detailing the yeast two-hybrid assay	**117**
Figure 7.10	The PrimPol C-terminus (370-560), encompassing the UL52 zinc finger, is responsible for auto-activation in the yeast two-hybrid assay	**117**

List of tables

Table 2.1	Primary antibodies	**42**
Table 2.2	Secondary antibodies	**42**
Table 2.3	Plasmid constructs for *T. brucei*	**43**
Table 2.4	Primers used to generate plasmid constructs for *T. brucei*	**43**
Table 2.5	Primers used to generate plasmid constructs for protein expression in *E. coli*	**46**
Table 2.6	Primers used to generate plasmid constructs for protein expression in *E. coli*	**46**
Table 2.7	Oligonucleotides used in this thesis for the primer extension assays	**49**
Table 2.8	Plasmid constructs for cultured human cells	**50**
Table 2.9	Primers used to generate plasmid constructs for cultured human cells	**50**
Table 2.10	Plasmid constructs for yeast two-hybrid	**57**
Table 2.11	Primers used to generate plasmid constructs for yeast two-hybrid analysis	**57**
Table 4.1	Summary of trans-lesion DNA synthesis activities of TbPrimPol1 and 2	**83**

Chapter 1

Introduction

1. Introduction

Genome stability is of upmost importance to life. DNA polymerases are essential for the duplication and maintenance of the genome but they cannot themselves begin synthesis of a DNA chain, and require the activity of specialised RNA polymerases called DNA primases. In eukaryotic cells distinct enzymes catalyse these two essential processes. This thesis contains the characterisation of a novel enzyme present in eukaryotic cells that is capable of both polymerase and primase activities, and aims to determine the role of this novel enzyme in human cells but also in an important protozoan pathogen. The introduction of this thesis will review our current understanding of DNA polymerases and DNA primases, the cellular processes of genome duplication and DNA damage tolerance of which DNA polymerases are central, and also introduce the African trypanosome and the novel primase-polymerase CCDC111.

1.1. DNA Polymerases

In 1956 an enzyme was discovered in *Escherichia coli* that could catalyse the accurate replication of deoxyribonucleic acid (DNA), and the name “DNA Polymerase” was coined (Kornberg *et al*., 1956; Lehman *et al*., 1958). Shortly after, an enzyme with identical activity was partially isolated from mammalian tissues (Bollum and Potter, 1958). Over 50 years later, a plethora of DNA polymerases have been identified in all branches of life. These proteins are essential for genome duplication and are critical in protecting cells against the affects of DNA damage.

1.1.1. Overview of eukaryotic DNA polymerases

DNA polymerases are enzymes that synthesise DNA, and are required for all DNA synthetic processes that occur within a cell. The human genome encodes at least 15 distinct DNA-dependant DNA polymerases (Figure 1.1). They can be broadly classified into four families based on their phylogenetic relationships with other DNA polymerases, the A-, B-, X-, and Y-families (Ito and Braithwaite, 1991 and 1993; Ohmori *et al*., 2001). The faithful duplication of the nuclear genome in eukaryotes is largely carried out by Pol alpha (α), delta (δ) and epsilon (ε), all B-family polymerases (reviewed in Garg and Burgers, 2005a), due of their phylogenetic relationship with Pol II from *E. coli*. The replication of the mitochondrial genome is carried out by Pol gamma (γ) (reviewed in Kaguni, 2004), an A-family polymerase, owing to its phylogenetic relationship to Pol I from *E. coli*. As the sole mitochondrial DNA polymerase, Pol γ is also required for DNA repair processes that occur within this essential organelle. The repair of the nuclear genome is largely performed by the X-family polymerases (reviewed in Uchiyama *et al*.,

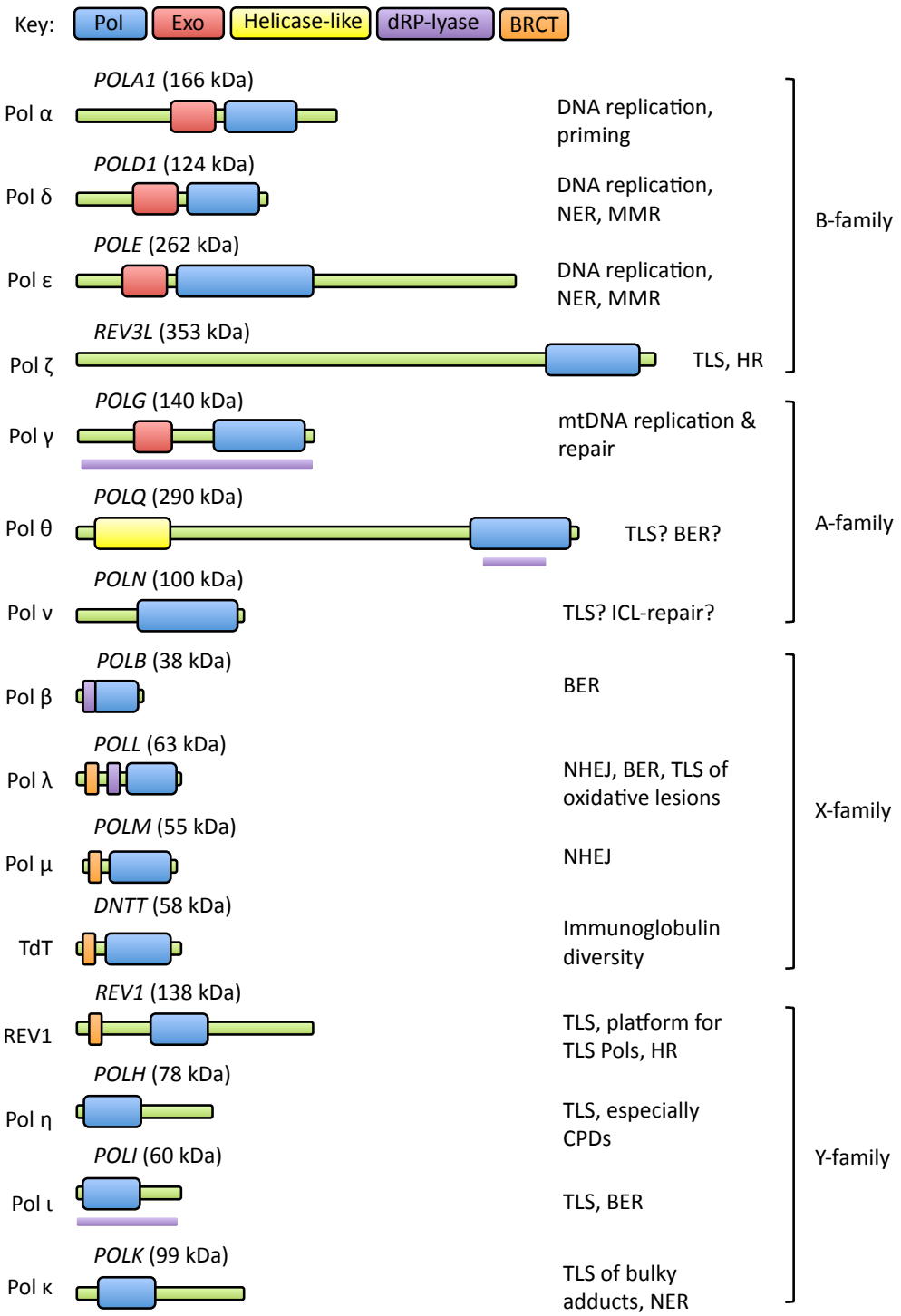


Figure 1.1. The human genome encodes 15 DNA-dependant DNA polymerases.
15 DNA-dependant DNA polymerases have been characterised to date in human cells, each named with a Greek letter. The domain structure of each polymerase catalytic subunit is shown: blue – polymerase domain, red – 3'-5' exonuclease domain, yellow – helicase-like domain, purple – 5'-deoxyribose phosphate (dRP)-lyase domain (purple line – dRP-lyase activity), orange – BRCA1 C-terminal (BRCT) domain. Although Pol α has a 3'-5' exonuclease domain it is inactive. Gene name and molecular weight of catalytic subunit is shown above each schematic and reported cellular function(s) of each polymerase shown alongside. BER – base excision repair, NER – nucleotide excision repair, MMR – mismatch repair, TLS – trans-lesion synthesis, NHEJ – nonhomologous end-joining, HR – homologous recombination, ICL – interstrand crosslink. The human DNA polymerases can be divided into four families: the B-, A-, X-, and Y-families, as indicated. Figure adapted from Lange *et al*., 2011.

2009), Pol beta (β), lambda (λ), and mu (μ), classified due to their homology with human Pol β. These enzymes are dedicated to repairing single-base lesions and double-strand breaks in DNA. An X-family member exempt from the polymerase nomenclature is Terminal Transferase (TdT) that plays a role in immunoglobulin diversity. Despite the numerous DNA repair mechanisms present within a cell, some DNA damage will inevitably escape repair and be encountered by the DNA replication machinery, resulting in DNA replication stalling. One solution is provided in the form of the Y-family polymerases, eta (η), iota (ι), kappa (κ), and REV1 (reviewed in Sale *et al*., 2012), classified due to their homology with *E. coli* DinB. These specialised polymerases are capable of synthesising DNA opposite DNA lesions that would otherwise stall a replicative DNA polymerase, thus allowing completion of replication in the presence of DNA damage. The B-family polymerase Pol ζ is also required for this process (reviewed in Gan *et al*., 2008). In the nucleus there are also two A-family polymerases, theta (θ) and nu (ν), whose cellular roles remain unclear.

1.1.2. Mechanism and structure of a DNA polymerase

Our understanding of the chemistry of DNA synthesis originates from Arthur Kornberg's and colleagues discovery of the first DNA polymerase in *E. coli* (now known as Pol I) (Kornberg *et al*., 1956; Lehman *et al*., 1958). The mechanism outlined in these seminal studies, detailed in Figure 1.2, has proven true for all DNA and RNA polymerases identified to date (Rothwell and Waksman, 2005). A DNA polymerase uses single-stranded DNA as a template to assemble an exact complementary replica by successive polymerisation of four complementary deoxynucleotides. However, a universal feature of DNA polymerases is their inability to initiate DNA synthesis *de novo*, i.e. starting with the polymerisation of two deoxynucleotides. Rather, a DNA polymerase extends from the 3' end of an already existing DNA or RNA chain (called a primer), synthesising DNA in a 5' to 3' direction. The mechanism of DNA synthesis is a repetitive cycle; first, the incoming deoxynucleotide 5' triphosphate (dNTP) substrate is paired by hydrogen bonding to the complementary templated base. The polymerase then catalyses the nucleophillic attack of the 3' hydroxyl group of the primer terminus on the α-phosphate group of the dNTP to be added. A phosphoryl transfer reaction occurs producing a phosphdiester bond that is the backbone of the DNA molecule. As with all enzymes that use dNTPs as substrates, metal ions are essential for catalysis, with magnesium ions (Mg^{2+}) presumably used *in vivo*. Two of the 3 phosphates present on the dNTP are subsequently released as pyrophosphate (PPi), and the catalytic cycle begins again at the next templated base if necessary.

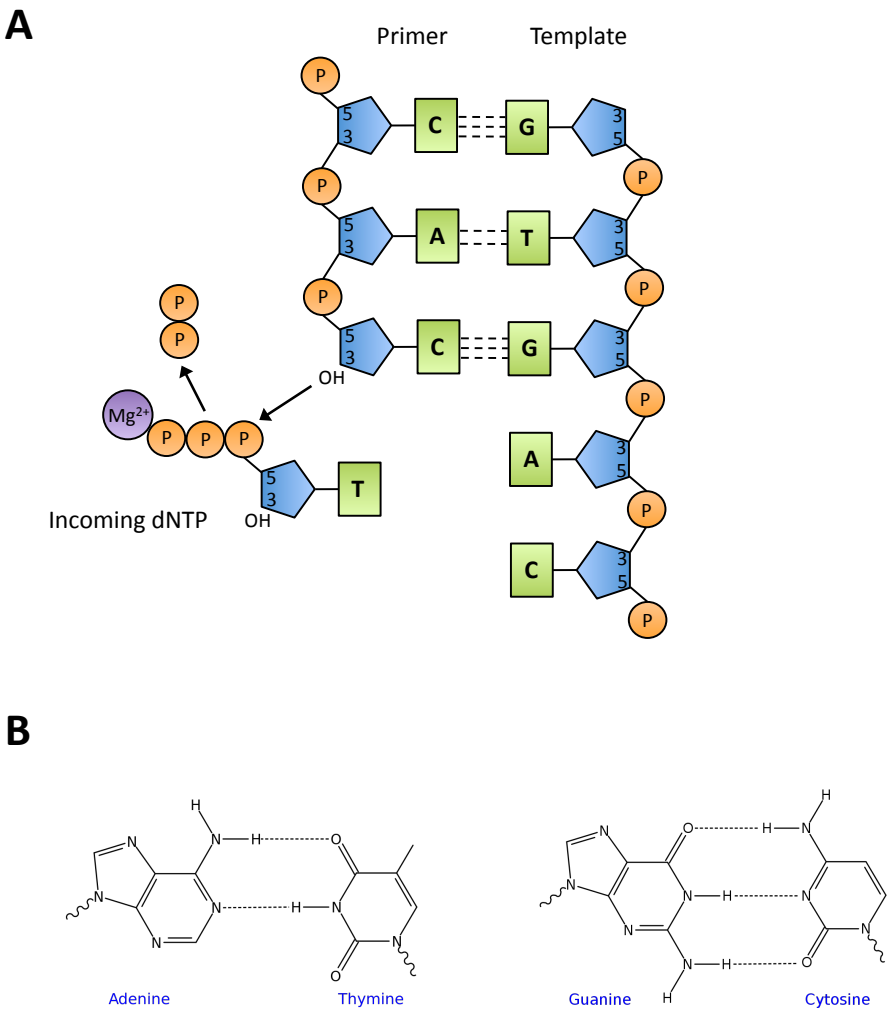


Figure 1.2. The enzymatic synthesis of DNA.
(**A**) A simplified DNA molecule is depicted, composed of a phosphate (orange circles) deoxyribose sugar (blue pentagons) backbone and four different nucleobases (green rectangles): guanine (G), adenine (A), thymine (T), and cytosine (C). The 3rd and 5th carbons of the deoxyribose sugars are indicated. Black lines indicate covalent linkages and dashed lines hydrogen bonds. A DNA polymerase catalyses the nucleophillic attack of the 3' hydroxyl (OH) group of the primer terminus on the α-phosphorus of the incoming complementary 5' deoxynucleotide triphosphate (dNTP). This results in a phosphoryl transfer reaction which adds a 5' deoxynucleotide monophosphate to the primer terminus and releases pyrophosphate (PPi). Two divalent cations, presumably magnesium (Mg^{2+}, purple circle) *in vivo*, are essential for DNA polymerase catalysis, and can be chelated by the incoming dNTPs. Adapted from Kornberg, 1959. (**B**) Chemical structure of the nucleobases of DNA in their Watson-Crick pairs, the non-covalent hydrogen bonds between the pairs are shown as dashed lines.

Understanding of the molecular basis of DNA synthesis began with the first DNA polymerase structure to be solved crystallographically. This was of the so-called Klenow fragment of *E. coli* DNA Polymerase I, which encompasses the polymerase and exonuclease domain (Klenow and Overgaard-Hansen, 1970). The structure of the DNA polymerase domain was likened to a human right hand, composed of a "palm", "fingers", and "thumb" sub-domains that together form a "U" shaped cleft (Figure 1.3) (Ollis *et al*., 1985). Despite the amino acid sequence diversity among DNA polymerases throughout all branches of life, the overall architecture of the active site proved to be similar for all DNA polymerase structures solved to date (Joyce and Steitz, 1994; Rothwell and Waksman, 2005). The DNA molecule resides in the palm of the polymerase and is grasped by the fingers and thumb sub-domains (Figure 1.3c). The fingers sub-domain makes important interaction with the incoming dNTP substrate and the paired template base, and undergoes a conformational change at each nucleotide addition step. The palm sub-domain contains the catalytic carboxylates required for phosphoryl transfer to create the phosphodieser bond to link the deoxynucleotides. Lastly, the thumb sub-domain interacts with and thereby positions the double-stranded DNA. Structural studies of DNA polymerases also delineated the importance of metal ions in catalysis. Phosphorly transfer is catalysed by a two-metal ion mechanism, first proposed by analogy to the 3'-5' exonuclease reaction (Beese and Steitz, 1991; Steitz 1993). Both metal ions are present in the active site; one interacts with the three phosphates of the incoming dNTP (metal ion B) whilst the second interacts with the hydroxyl group and the α-phosphate of the dNTP (metal ion A) (Figure 1.4). Both metal ions A and B stabilise the transition state of the enzyme. Metal ion A activates the hydroxyl group and facilitates the nucleophilic attack required for phosphoryl transfer, whilst metal ion B facilitates the leaving of PPi.

1.1.3. Processivity and fidelity – important features of a DNA polymerase

Processivity is an important attribute of a DNA polymerase, defined as the number of deoxynucleotides incorporated by the DNA polymerase into a DNA chain in a single binding event to the DNA template (Von Hippel *et al*., 1994; Bambara *et al*., 1995). This can vary substantially depending on the function of the DNA polymerase. For instance, the B-family polymerase Pol δ from *Saccharomyces cerevisiae* is able to polymerise over 7000 deoxynucleotides in a single binding event to a DNA template *in vitro* (Burgers, 1991). This high processivity is ideal for the timely duplication of the nuclear genome in eukaryotic cells. In contrast, the Y-family polymerase Pol η from *S. cerevisiae*, incorporates only 6 or 7 deoxynucleotides into a DNA chain in a single binding event *in vitro* (Washington *et al*., 1999). This low processivity, also referred to as distributive DNA synthesis, is ideal as this polymerase is required to synthesise DNA opposite DNA

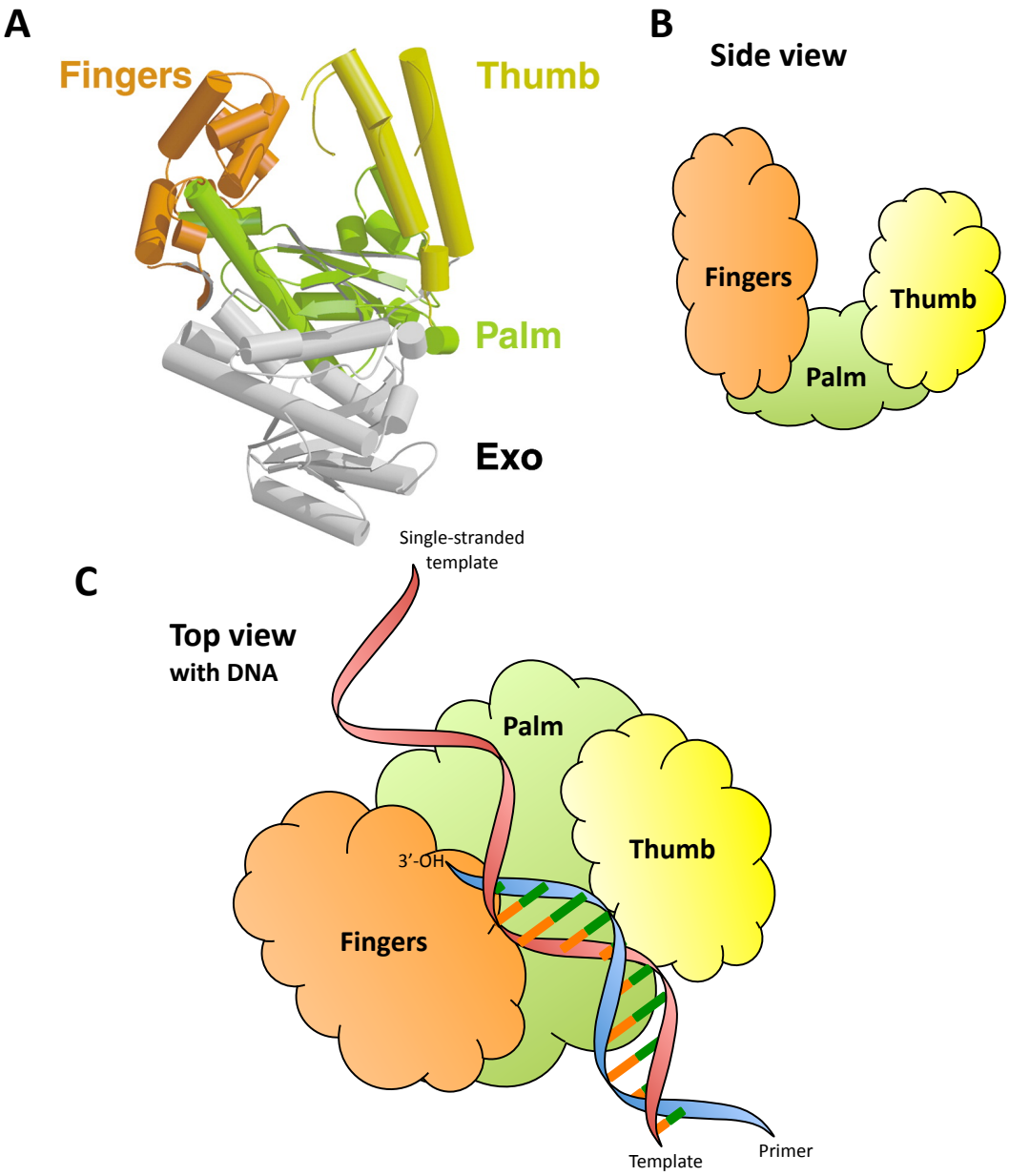


Figure 1.3. Structure of a DNA polymerase.
(**A**) The crystallographic structure of the *E. coli* DNA Pol I Klenow fragment, which encompasses the polymerase and exonuclease domain, was the first crystallographic structure solved of a DNA polymerase (Ollis *et al*., 1985). The structure was likened to a human right hand composed of palm (green), fingers (orange), and thumb (yellow) sub-domains, together forming a "U" shaped cleft. Figure taken from Beard and Wilson, 2001. (**B**) Cartoon depiction of a polymerase domain. Despite amino acid sequence divergence between DNA polymerases, all structures of DNA polymerases solved to date have followed this human right hand architecture. (**C**) Cartoon depiction of a polymerase with DNA. The DNA molecule sits in the "U" shaped cleft gripped by the fingers and thumb sub-domains. Prior to polymerisation the fingers domain is in an 'open' conformation and interacts with the incoming complementary dNTP, and then changes to a 'closed' conformation upon dNTP pairing with the templated base. The palm domain contains the catalytic carboxylate residues essential for phoshoryl transfer between the dNTP and the 3' hydroxyl (OH) (indicated) of the primer terminus. The thumb sub-domain interacts with the double-stranded DNA

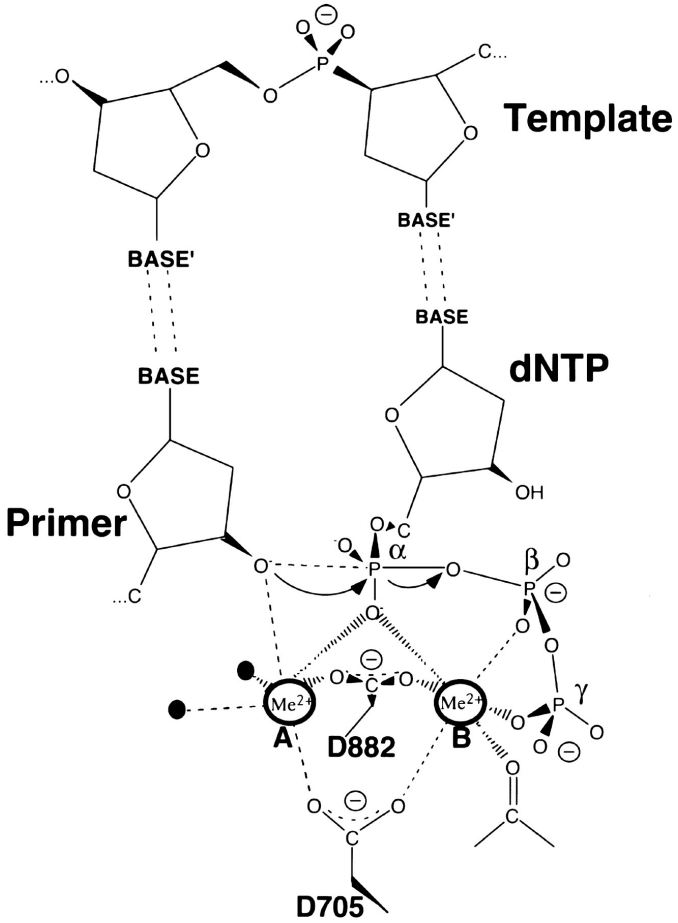


Figure 1.4. The two-metal ion mechanism of DNA polymerases.
The phosphoryl transfer reaction of all DNA polymerases is catalysed by a two-metal ion mechanism, originally proposed by analogy to the well characterised catalysis mechanism of 3' exonucleases (Beese and Steitz, 1991; Steitz, 1993). The two conserved aspartates depicted are numbered according to *E. coli* DNA Pol I. The active site contains two divalent metal ions (Me^{2+}), A and B, that are coordinated by the two carboxyaltes of the catalytic aspartates. Both metal ions stabilise the pentacovalent transition state formed during catalysis. Metal ion A activates the 3' hydroxyl (OH) group for nucleophillic attack of the α-phosphate of the incoming deoxynucleotide 5' triphosphate (dNTP). Metal ion B interacts with the β and γ phosphates of the dNTP and facilitates their leaving. Two water molecules (black circles) are bound to metal ion A. This figure is taken from Brautigam and Steitz, 1998.

lesions that only span a few deoxynucleotides of a DNA chain. Processivity of a DNA polymerase is largely determined by interactions between the thumb sub-domain of the polymerase and the primer-template DNA, illustrated by B-family polymerases having large thumb sub-domains whilst Y-family polymerases have stubby thumb sub-domains (Rothwell and Waksman, 2005; Sale *et al*, 2012). Additional proteins that interact with the DNA polymerase can also greatly influence the processivity of DNA synthesis through two different mechanisms (reviewed in Kelman *et al*., 1998). The first mechanism is exemplified by so-called sliding clamps, such as the eukaryotic Proliferating Cell Nuclear Antigen (PCNA), which encircles the DNA molecule and physically clamps the DNA polymerase to the template thereby preventing dissociation of the polymerase (Krishna *et al*., 1994). The second mechanism is through interaction of the DNA polymerase with a specific accessory subunit that results in increased fidelity, exemplified by the accessory subunit of the mitochondrial replicase Pol γ (Lim *et al*., 1999).

Fidelity of DNA synthesis is also an important attribute of a DNA polymerase, and again, varies depending on the function of the polymerase. High fidelity DNA synthesis is extremely important as accurate duplication of DNA is essential for life, whilst low fidelity DNA synthesis is important for generating genetic diversity in the development of the immune system (reviewed in Weill and Reynaud, 2008). Polymerase fidelity is defined as the ratio of incorporation of the correct over incorrect nucleotide, which is determined largely by the efficiency of the polymerase to insert the correct nucleotide (Beard *et al*., 2002). The fidelity of DNA replication in human cells is astonishingly high with a single error occurring for every 10^9-10^{10} bases replicated (Loeb, 1991), but this is only partly attributed to the accuracy of DNA polymerases, as post-replication mismatch repair improves the fidelity of DNA replication by several orders of magnitude (reviewed in Jiricny, 2006). High fidelity of DNA synthesis by a DNA polymerase is achieved by a number of mechanisms (reviewed in Kool, 2002). Watson-Crick hydrogen bonding only contributes in part to a DNA polymerases ability to select the correct over incorrect dNTP (Loeb *et al*., 1974), and it is in fact the active site geometry of the polymerase that is proposed to play large part in maintaining fidelity (Goodman, 1997). The active site of a DNA polymerase fits snugly around bases in a Watson-Crick pair, whilst mismatched bases result in steric clashes (Johnson and Beese, 2004). These clashes are predicted to reduce the binding affinity of the dNTP, affect the conformational change required to set up the active site geometry for catalysis, and reduce the rate of phosphodiester bond formation, thus making the selection of the correct dNTP much more favourable (McCulloch and Kunkel, 2008). To further increase the fidelity of DNA synthesis, some family-B DNA polymerases also contain a 3'-5' exonuclease domain allowing

proofreading of newly synthesised DNA. In the event of polymerising a mismatched base, the subsequent extension is of lower efficiency and so is delayed, allowing the primer terminus to enter the exonuclease active site where the mismatched base is excised allowing the correct nucleotide to be incorporated (reviewed in Steitz, 1999).

1.2. Genome Duplication

The quintessential role of DNA polymerases is genome duplication. Complete and accurate DNA replication prior to cell division is essential for maintaining genome stability in all organisms. This process occurs during a defined period in the cell cycle, termed the DNA synthesis (S)-phase, and is highly regulated to ensure coordination with the cell cycle processes of mitosis and cytokinesis, in addition to coupling with DNA repair, chromatin re-assembly and retention of epigenetic markers, and transcription. It is the family-B polymerases α, δ, and ε that are central to DNA replication.

1.2.1. Replication origins – where to build a replication fork

A pioneering study by Huberman and Riggs in the 1960's noted that the total chromosomal DNA in a mammalian cell was of a tremendous length, and in order to be replicated completely during a single S-phase, DNA replication must initiate from multiple sites throughout the genome (Figure 1.5) (Huberman and Riggs, 1966). They later referred to these sites as "origins" and found that they were between 15 and 300 kilo bases apart (Huberman and Riggs, 1968). We now understand that the human genome contains an excess of origins (10,000 and 100,000), which have no sequence consensus, and only a subset are fired once per cell cycle, during S-phase, to allow complete genome duplication (reviewed in Méchali, 2010). It is at origins that replication forks assemble, the structures by which cellular DNA is replicated, first visualised in *E. coli* (Cairns, 1963), and shortly after by Huberman and Riggs, with the latter initiating bi-directional DNA synthesis (Huberman and Riggs, 1968). In order to assemble replication forks at an origin it must first be licensed (Figure 1.6), a process that occurs during late mitosis and G1 phase in which a pre-replication complex (pre-RC) assembles that can recruit the necessary components for the initiation of DNA replication. The first component of the pre-RC is the origin recognition complex (ORC) (Bell and Stillman, 1992), which together with two other proteins (CDC6 and CDT1) loads the heterohexameric mini-chromosome maintenance (MCM) 2-7 complex, the core of the putative replicative DNA helicase. However the MCM helicase is loaded in its inactive form, and at the G1-S transition the activities of two cell cycle kinases facilitate the loading of replication proteins onto the pre-RC. These activate the MCM helicase and

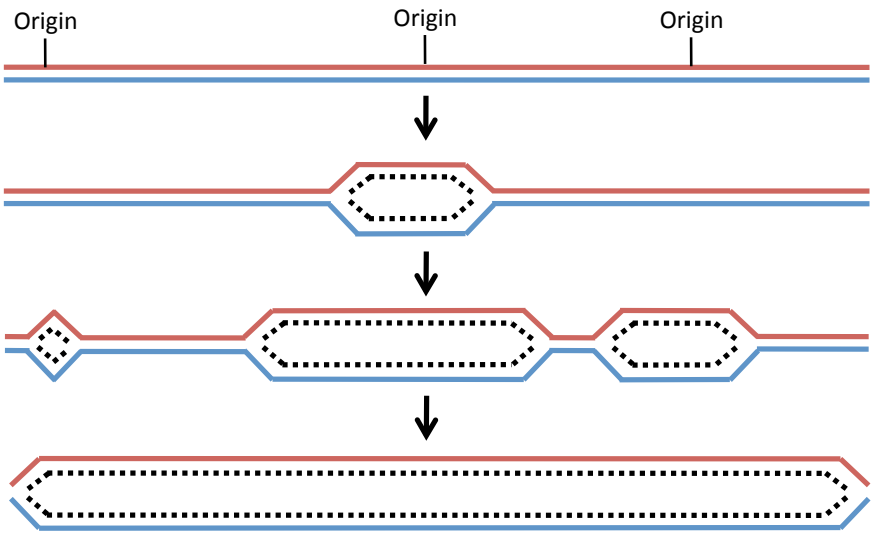


Figure 1.5. The mechanism of chromosomal DNA replication.
The DNA double helix is depicted as a pair of horizontal lines, coloured lines represent template DNA and dashed lines represent nascent DNA. Chromosomal DNA replication in eukaryotic cells is initiated at multiple sites throughout the genome, called replication origins (indicated). DNA synthesis initiates stochastically throughout the genome using only a subset of the available replication origins, and initiation is bidirectional with each origin producing two replication forks which progress in opposite directions to duplicate the genome. Figure adapted from Huberman and Riggs, 1968.

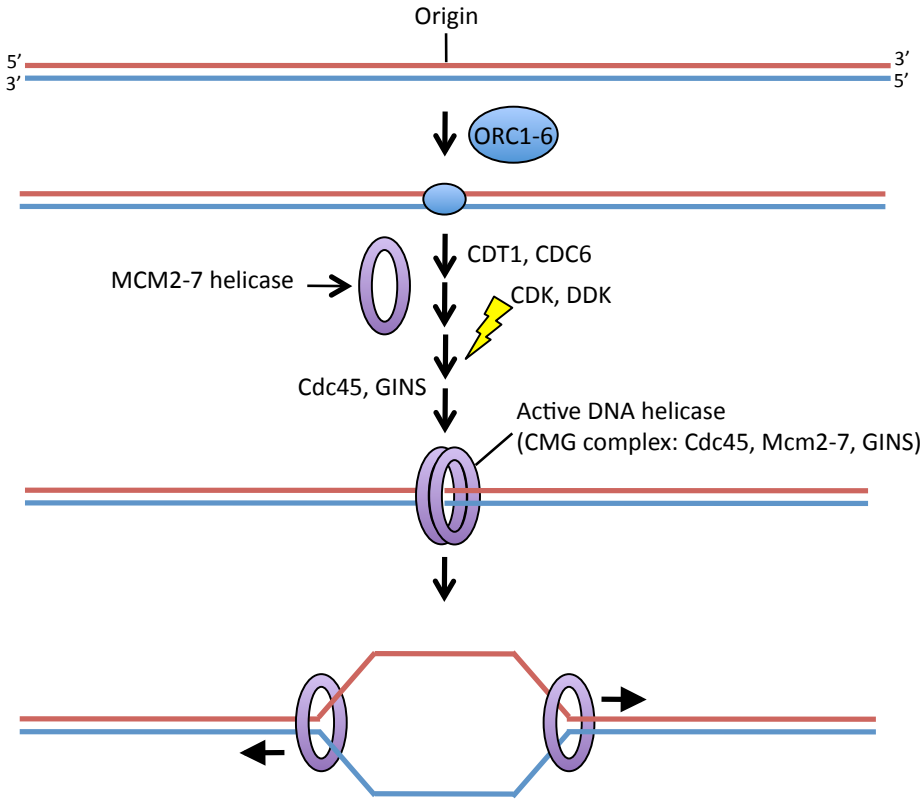


Figure 1.6. The licensing and firing of replication origins.
The assembly of a pre-replication (RC) complex at a replication origin and subsequent origin firing is depicted. The heterohexameric origin recognition complex (ORC1-6) directly binds origins of replication, which in higher eukaryotes have no sequence consensus. CDC6 and CDT1 are then recruited and are responsible for the recruitment and loading of the heterohexameric ATP-dependant DNA helicase mini-chromosome maintenance complex (MCM2-7). When loaded the MCM2-7 helicase is in its inactive form and marks the completion of the pre-RC and origin licensing. The MCM2-7 helicase is activated by the binding of CDC45 and the GINS complex and the concerted actions of S-phase cyclin-dependant kinases (S-CDKs) and DBF4-dependant kinase (DDK), thus allowing unwinding of duplex DNA at the replication origin and the beginning of DNA replication. For simplicity, a number of factors are omitted from the figure, including Pol ε, which actually loads prior to origin firing, suggesting a role in assembly of the replication fork.

thus allow unwinding of the duplex DNA at the origin and assembly of the replication fork (reviewed in Bell and Dutta, 2002).

1.2.2. Initiation – biogenesis of a replication fork

The general mechanism of chromosomal DNA replication in eukaryotes was first characterised by biochemical studies using the Simian Virus 40 (SV40) replication system with cell extracts and purified mammalian proteins (reviewed in Waga and Stillman, 1998). Indeed, essential DNA replication components were originally identified in these studies, such as the heterotrimeric single-stranded DNA binding protein Replication Protein A (RPA) (Wobbe *et al*., 1987; Fairman and Stillman 1988; Wold and Kelly 1988) and the DNA polymerase clamp loading complex Replication Factor C (RFC) (Tsurimoto and Stillman, 1989). Similarly these studies uncovered the essential role of the DNA polymerase sliding clamp PCNA in replication (Prelich *et al*., 1987) and led to the original purification of human Pol δ (Melendey and Stillman, 1991).

The first DNA polymerase activity required at the unwound origin is that of Pol α. In all eukaryotic organisms characterised to date, Pol α is a heterodimeric enzyme composed of a catalytic and accessory (B) subunit that exists associated to the heterodimeric DNA Primase, itself composed of a small catalytic (Prim1) and large accessory (Prim2) subunit (reviewed in Kaguni and Lehman, 1988). This complex, called Pol α-Prim, possesses the unique ability in eukaryotic cells of being able to initiate DNA synthesis *de novo*, due to the DNA-dependant RNA polymerase activity of DNA Primase (reviewed in Muzi-Falconi *et al*., 2003; Frick and Richardson, 2001; see section 1.6). Upon local unwinding of the duplex DNA at the origin, RPA first binds to and stabilises the single-stranded DNA (Figure 1.7). Pol α-Prim is then recruited to the single-stranded DNA and the primase component synthesises a short ~10 ribo-oligonucleotide molecule that provides the 3' hydroxyl group required to start DNA synthesis by Pol α, which subsequently extends the RNA primer with a DNA chain of ~20 nucleotides (Matsumoto *et al*., 1990; Murakami *et al*., 1992). Owing to the anti-parallel nature of duplex DNA and that DNA polymerases synthesise DNA in a 5' to 3' direction, the two strands of DNA are synthesised in a different manner during chromosomal DNA replication, which was first proposed by Okazaki and colleagues (Sakabe and Okazaki, 1966; Okazaki *et al*., 1967). The leading strand is synthesised continuously (or at least largely), whilst the lagging strand is synthesised in a discontinuous manner via ~200 nucleotide so-called Okazaki fragments. On the leading strand, Pol α-Prim is required to initiate DNA synthesis at replication origins, whilst on the lagging strand, it is required to initiate synthesis of each Okazaki fragment. Following synthesis of the RNA-DNA primer a DNA polymerase switch

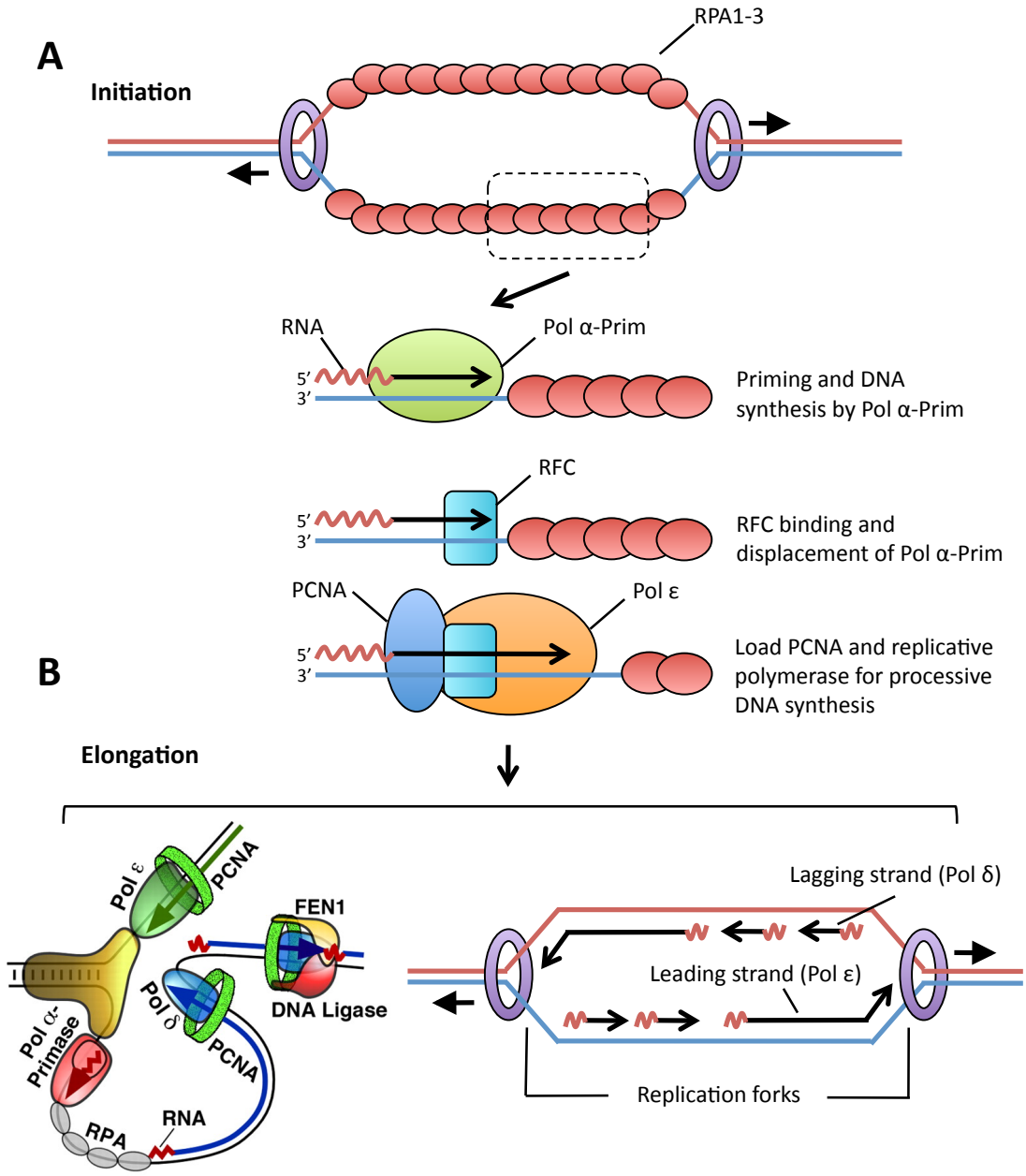


Figure 1.7. The initiation and elongation of DNA replication.
The initiation (**A**) and elongation (**B**) of DNA replication from a single replication origin is depicted. Unwinding of the DNA duplex by the active MCM2-7 helicase complex produces single-stranded DNA that is subsequently coated with the RPA heterotimer. The Pol α-Prim complex is recruited and using its DNA primase component synthesis a short RNA molecule *de novo,* providing the 3' hydroxyl group for Pol α to extend with a nascent DNA chain**.** Replication factor C (RFC) binds to the 3' terminus of the RNA-DNA primer and facilitates the dissociation of Pol α-Prim and the loading of the DNA polymerase processivity clamp PCNA. The replicative DNA polymerase is then loaded, Pol ε on the leading stand (shown above) and Pol δ of the lagging strand, and the elongation phase of DNA synthesis begins. The leading strand is synthesised largely continuously whilst the lagging strand is synthesised in discontinuous ~200 nucleotide Okazaki fragments. The division of labour at the replication fork by the replicative DNA polymerases is also depicted, taken from Burgers, 2009. A portion of this figure was adapted from Waga and Stillman, 1998.

occurs, displacing Pol α-Prim and loading the replicative DNA polymerases required for bulk DNA synthesis (Figure 1.7). This switch is mediated by RFC-dependant loading of the homotrimeric polymerase clamp PCNA (Tsurimoto and Stillman, 1991a and b), and marks the beginning of the elongation phase of DNA replication.

1.2.3. Elongation – progression of the replication fork

Using the cell-free SV40 replication system and purified mammalian proteins both leading and lagging strand synthesis were reconstituted *in vitro* (Tsurimoto *et al*., 1990; Waga and Stillman, 1994). However one caveat to this system was its requirement for only two DNA polymerases, Pol α and δ, whilst genetic studies in yeast demonstrated that a third B-family DNA polymerase, Pol ε, was also required for chromosomal DNA replication (Morrison *et al*., 1990). Studies from the Kunkel laboratory have largely solved this discrepancy using active site mutants of Pol δ and ε with normal rates of DNA synthesis but high error frequencies. They demonstrated in *S. cerevisiae* that Pol ε participates in leading strand synthesis (Pursell *et al*., 2007) and that Pol δ is responsible for the majority of lagging strand synthesis (Nick McElhinny *et al*, 2008). This division of labour at the replication fork, depicted in Figure 1.7b, was shown to be evolutionarily conserved in *S. pombe* also (Miyabe *et al*., 2011), and given the conservation of replication fork components from yeast to man (Errico and Costanzo, 2010), a similar division of labour would be expected at the human replication fork, although this is yet to be tested. One apparent area of controversy remaining is the extent of Pol ε-dependant DNA synthesis on the leading strand, with an alternate model existing in which Pol ε is required for only initial DNA synthesis on the leading strand, with Pol δ then taking over (reviewed in Pavlov and Shcherbakova, 2009). Pol δ and ε are both mulit-subunit enzymes of high processivity and with a 3'-5' exonuclease proofreading activity ensuring high fidelity (reviewed in Kunkel and Burgers, 2008). On the lagging strand Pol δ synthesises each Okazaki fragment up until the RNA-DNA primer of the next Okazaki fragment, then the process of Okazaki fragment maturation occurs with the combined activities of the flap endonuclease FEN1 and DNA Ligase I (reviewed in Waga and Stillman, 1998; Burgers, 2009). On the leading strand, DNA synthesis is largely thought to be continuous in the absence of DNA damage, and so the leading strand polymerase will synthesise DNA until the replication fork converges with an adjacent replicon. The termination of DNA replication is less well understood than initiation and elongation. Evidence from SV40 and yeast studies indicate that sister chromatids become catanated at replication termination sites and require DNA topoisomerase II for resolution (DiNardo *et al*., 1984; Fields-Berry and DePamphilis *et al*., 1989; Sundin and Varshavsky, 1980; Cuvier *et al*., 2008). It has been suggested that, in contrast to DNA replication initiation,

DNA replication termination occurs randomly (Santamaria *et al*., 2000). However, a recent study in *S. cerevisiae* has shown that eukaryotic chromosomes do contain DNA replication termination regions, like in bacteria, and these regions contain replication fork pausing elements and mediate replication fork fusion (Fachinetti *et al*., 2010).

1.3. Overview of DNA damage and repair

DNA is subject to continuous damage within a cell, which if not repaired can have catastrophic consequences, particularly if encountered by the replication fork. Lesions within DNA arise from three main sources (Figure 1.8). Firstly, the DNA molecule has a limited chemical stability and as a result some of the chemical bonds within DNA have a tendency to spontaneously decay (reviewed in Lindahl, 1993). The spontaneous hydrolysis of the base-sugar bond in DNA results in the removal of the base residue whilst leaving the phosphodiester backbone intact, creating an abasic site, which can impede DNA replication. Additionally, hydrolytic deamination of the bases themselves can, for example, change a cytosine to a uracil, which will be misread by the replicative DNA polymerase creating a mutation. Secondly, (by)-products of normal cellular metabolism are an abundant source of DNA damage. Reactive oxygen species (ROS) such as superoxide anions, hydrogen peroxide, and hydroxyl radicals, created from oxidative phosphorylation in mitochondria, readily attack the DNA molecule (reviewed in Cadet *et al*., 2003). Perhaps the most abundant DNA lesion in the cell results from the oxidation of guanine to produce the miscoding 8-oxo-2'-deoxyguanosine (8-oxo-G) lesion. Thirdly, environmental agents pose a significant threat to the genome, such as the ultraviolet (UV) component of sunlight, ionising radiation (IR), and genotoxic agents. Of particular interest to this thesis are the DNA lesions resulting from exposure of cells to UV radiation, detailed in Figure 1.9. UV irradiation of DNA results in the formation of covalent adducts between adjacent pyrimidines on the same DNA strand, predominantly cyclobutane pyrimidine dimers (CPDs) and pyrimidine (6-4) pyrimidone photoproducts (reviewed in Rastogi *et al*., 2010). It is UV-C radiation (< 280 nm) that is predominantly used in the laboratory and causes these lesions, however UV-B radiation (280-315 nm), which terrestrial organisms are exposed to, can also induce formation of these photoproducts. These lesions are bulky and distort the DNA double helix (Figure 1.9b) and therefore significantly block DNA transactions, such as replication and transcription, and therefore can prove lethal (reviewed in Batista *et al*., 2009).

Given the importance of maintaining genome stability it is not surprising that a plethora of distinct DNA repair pathways exist within cells, shown in Figure 1.8. The first observation of a DNA repair mechanism was over 60 years ago when two independent

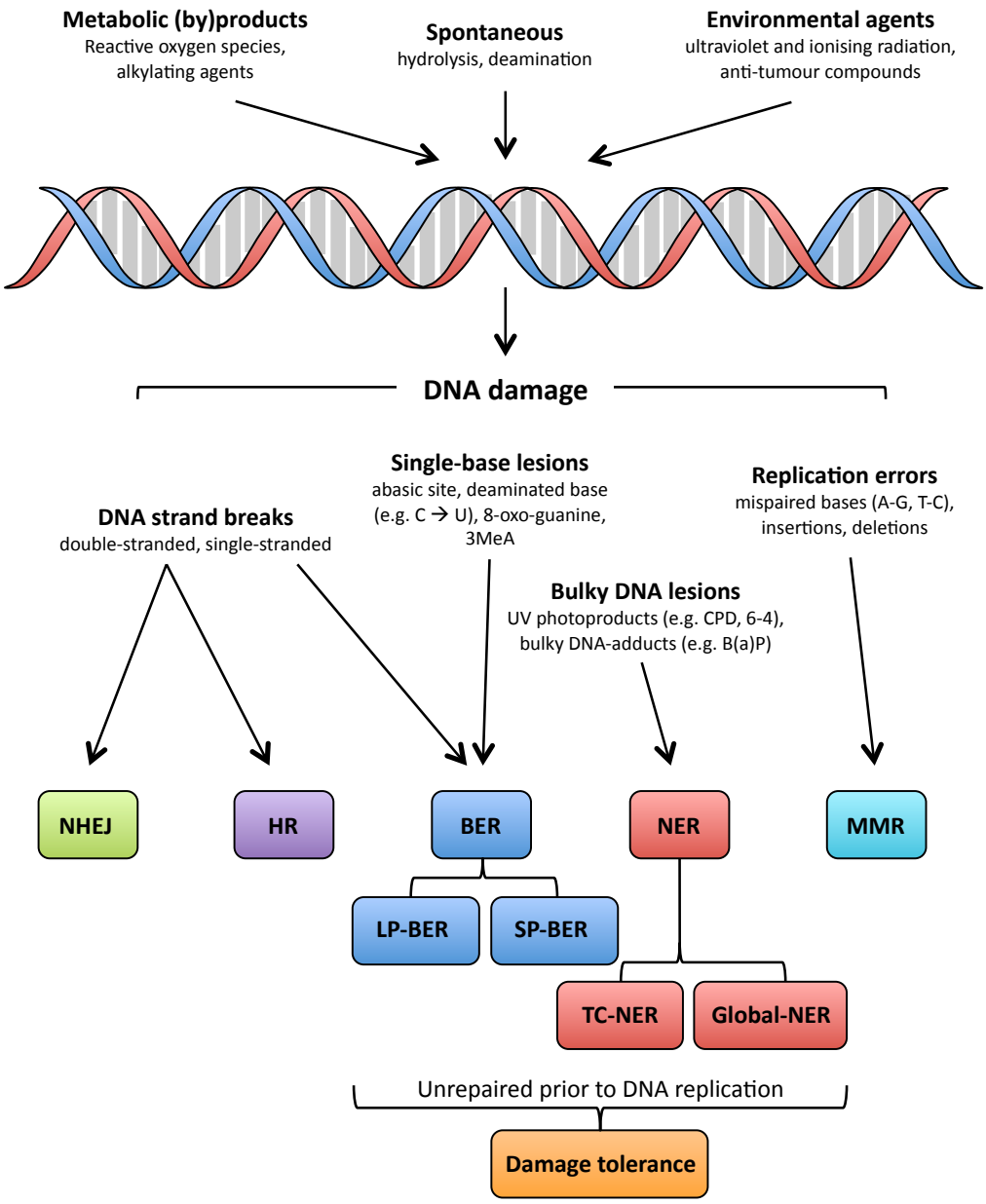


Figure 1.8. Overview of the causes and repair mechanisms of DNA damage in human cells. DNA damage is a fact of life. Depicted above are some of the numerous sources of DNA damage, the DNA modifications which they can produce, and the mechanisms dedicated to their repair (coloured boxes). The DNA molecule has an inherent instability as some of the chemical bonds within DNA are prone to decay. In addition, metabolic (by)products and environmental agents can produce a variety of DNA lesions. Double-strand breaks in DNA are repaired by nonhomologous end-joining (NHEJ) or homologous recombination (HR). Errors occurring during DNA replication, either due to misincorporation by the replicative polymerase or due to a mis-coding DNA lesion, are repaired by mismatch repair (MMR). Single-base lesions that cause little or no distortion of the DNA double helix and single-strand DNA breaks are repaired by base excision repair (BER), whilst bulky lesions that significantly distort the structure of DNA are repaired by nucleotide excision repair (NER). Even in the presence of fully functional BER and NER, the replication machinery will inevitably encounter DNA lesions, under these circumstances damage tolerance mechanisms are used to allow completion of DNA replication in spite of damage, and the damage can be subsequently repaired following replication fork progression.

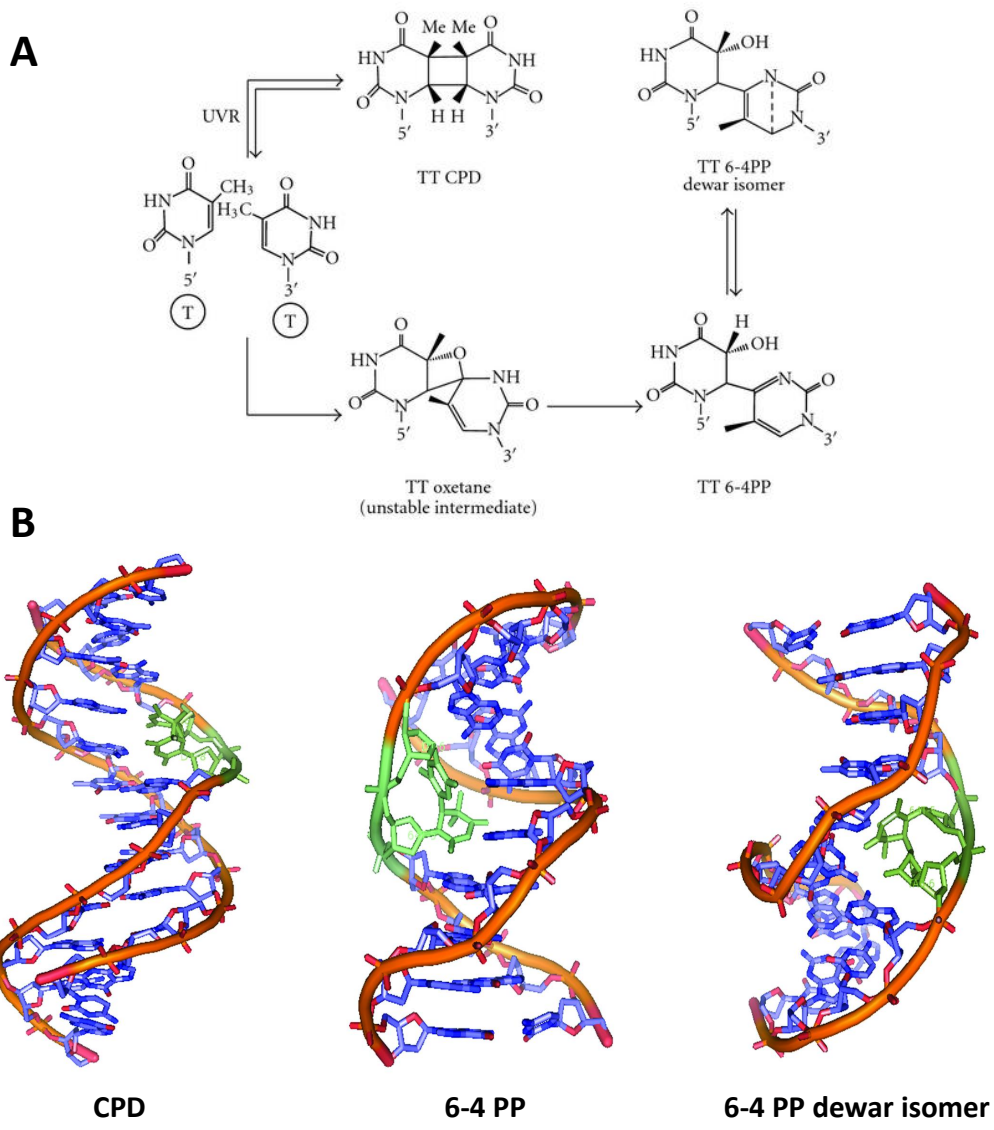


Figure 1.9. Exposure of DNA to UV radiation can result in the formation of photoproducts. UV irradiation of DNA can result in the crosslinking of adjacent pyrimidines on the same DNA strand, producing cyclobutane pyrimidine dimers (CPDs), pyrimidine (6-4) pyrimidone photoproducts (6-4 PPs), and a 6-4 PP dewar isomer. (**A**) Pathway of UV radiation (UVR) induced thymine-thymine (T-T) photoproducts. A CPD contains a four member ring structure involving C5 and C6 of neighbouring pyrimidines, whilst a 6-4 PP involves a linkage between C6 (5' T) and C4 (3' T). Exposure of a 6-4 PP to UV radiation can result in the formation of its valence dewar isomer and it's subsequent reversion. (**B**) Structures of DNA duplexes containing UV photoproducts (in green), a CPD (McAteer *et al*., 1998), 6-4 PP (Lee *et al*., 1999), and 6-4 PP dewar isomer (Lee *et al.*, 2000) are shown. Whilst a CPD only produces a subtle distortion of the DNA double helix, both the 6-4 PP and it's dewar isomer severely distort the DNA helix. It is because of this distortion that these lesions pose a substantial threat to completion of DNA replication and transcription. Figure taken from Rastogi *et al*., 2010.

researchers serendipitously observed the biological effects of UV radiation could be reversed by illumination with visible light (Kelner, 1949; Dulbecco, 1949). This phenomenon is now known as photoreactivation, an example of direct DNA repair, in which a light-dependant enzyme (photolyase) performs a series of photochemical reactions to reverse the dimerisation of adjacent pyrimidines caused by UV irradiation. Over a decade later thymine dimers were identified as stable, naturally occurring DNA lesions in cells exposed to UV light (Setlow *et al*., 1963), and the monitoring of these lesions within cells and their subsequent removal led to the discovery of excision repair in bacteria (Setlow and Carrier, 1964; Boyce and Howard-Flanders, 1964) and mammalian cells (Rasmussen and Painter, 1964). Over half a century later we can now broadly divide DNA repair into 6 distinct pathways: direct repair, mismatch repair (MMR), base excision repair (BER), nucleotide excision repair (NER), nonhomologous end-joining (NHEJ), and homologous recombination (HR) (Figure 1.8). All of these pathways, with the exception of direct repair, require a DNA synthetic step, highlighting the crucial role of DNA polymerases in protecting cells against the effects of DNA damage.

NER, BER, and MMR, all have a common general mechanism: first the damaged DNA is excised, a DNA polymerase then synthesises a replacement using the undamaged strand as a template, and finally the pieces are ligated together restoring the DNA double helix. NER is concerned with bulky, helix-distorting, DNA lesions that generally disrupt DNA transcription and replication, such as the UV-induced pyrimidine dimers (Figure 1.9). In humans NER requires at least 30 different proteins that are needed to make excisions either side of the damaged DNA, typically excising 30 nucleotides. The replicative polymerases δ and ε, along with the Y-family Pol κ, are then required for the DNA synthesis step (reviewed in Lehmann, 2011). Damaged bases that are not recognised by NER will be typically repaired by BER, which is concerned with small, single-base lesions, which do not distort the DNA helix but frequently miscode. The X-family Pol β is a central player in BER. MMR is concerned with correcting the replication errors caused during DNA replication and it is the replicative polymerases that are involved. NHEJ and HR are both concerned with repairing DNA breaks that occur in both strands of the DNA molecule. NHEJ is predominantly used during G1 phase due to the lack of a sister chromatid, and is inherently an error-prone process, whilst HR is used when a sister chromatid is present following genome duplication and is inherently error-free. The X-family polymerases are required for NHEJ, whilst the both B-family and Y-family polymerases have been implicated in HR (Lange *et al*., 2011).

1.4. DNA damage tolerance

Replication of the 6 billion nucleotides that constitute the human genome is a herculean task that is further complicated by the presence of DNA damage. DNA can become distorted and its bases modified by various environmental insults and endogenous processes, and if not corrected prior to replication, can result in the physical blockage of the replicative DNA polymerases. This is because the high fidelity of the replicative polymerases is not compatible with replicating damaged DNA. If not resolved, stalled replication forks pose a severe threat to genomic stability. The DNA lesion can no longer be repaired by BER or NER as the lesion is in single-stranded DNA. Reversal of the replication fork, thus placing the lesion in double-stranded DNA, has been shown in bacteria (Courcelle *et al*., 2003), but little evidence of this process exists in eukaryotes (Lopes *et al*., 2006). One solution that is present in all branches of life is DNA damage tolerance in which DNA replication proceeds in spite of the damaged DNA. DNA damage tolerance can be broadly divided into two pathways: error-free template switching and "error-prone" trans-lesion synthesis (TLS), which are both depicted in Figure 1.10.

1.4.2. Discovery of DNA damage tolerance

The study of DNA damage tolerance began with the seminal work of Rupp and Howard-Flanders in 1968. They observed that in UV irradiated *E. coli*, which were deficient in the UV photoproduct repair pathway NER, DNA was synthesised with only a minimal delay, however the daughter strands were synthesised discontinuously. These discontinuities or gaps were of a size similar to the spacing of pyrimidine dimers within the template strand, suggesting they were present opposite the lesions, and these gaps were subsequently sealed (Rupp and Howard-Flanders, 1968; reviewed in Bridges, 2005). This phenomenon was termed "post-replication repair", in reference to the repair of the post-replication gaps rather than the repair of the UV photoproducts, and was later demonstrated in mammalian cells (Lehmann, 1972). Shortly after, post-replication repair became clinically relevant with the discovery of a defect in this process in a group of patients with xeroderma pigmentosum (XP) (Lehmann *et al*., 1975), an inherited disorder characterised by hypersensitivity of the skin to sunlight. Meanwhile genetic studies in budding yeast had identified numerous genes involved in the cellular response to UV radiation: the *RAD* genes (Cox and Parry, 1968) of which mutants were sensitive to UV and/or IR, and the *REV* genes (Lemontt, 1971) of which mutants had reduced UV-induced mutagenesis; and it was these genes that were required for post-replication repair (DNA damage tolerance) in budding yeast (Di Caprio and Cox, 1981; Prakash, 1981). These genes were subsequently classified as the *RAD6* epitasis group, with both Rad6 and Rad18 being essential for damage tolerance, and could be divided into two

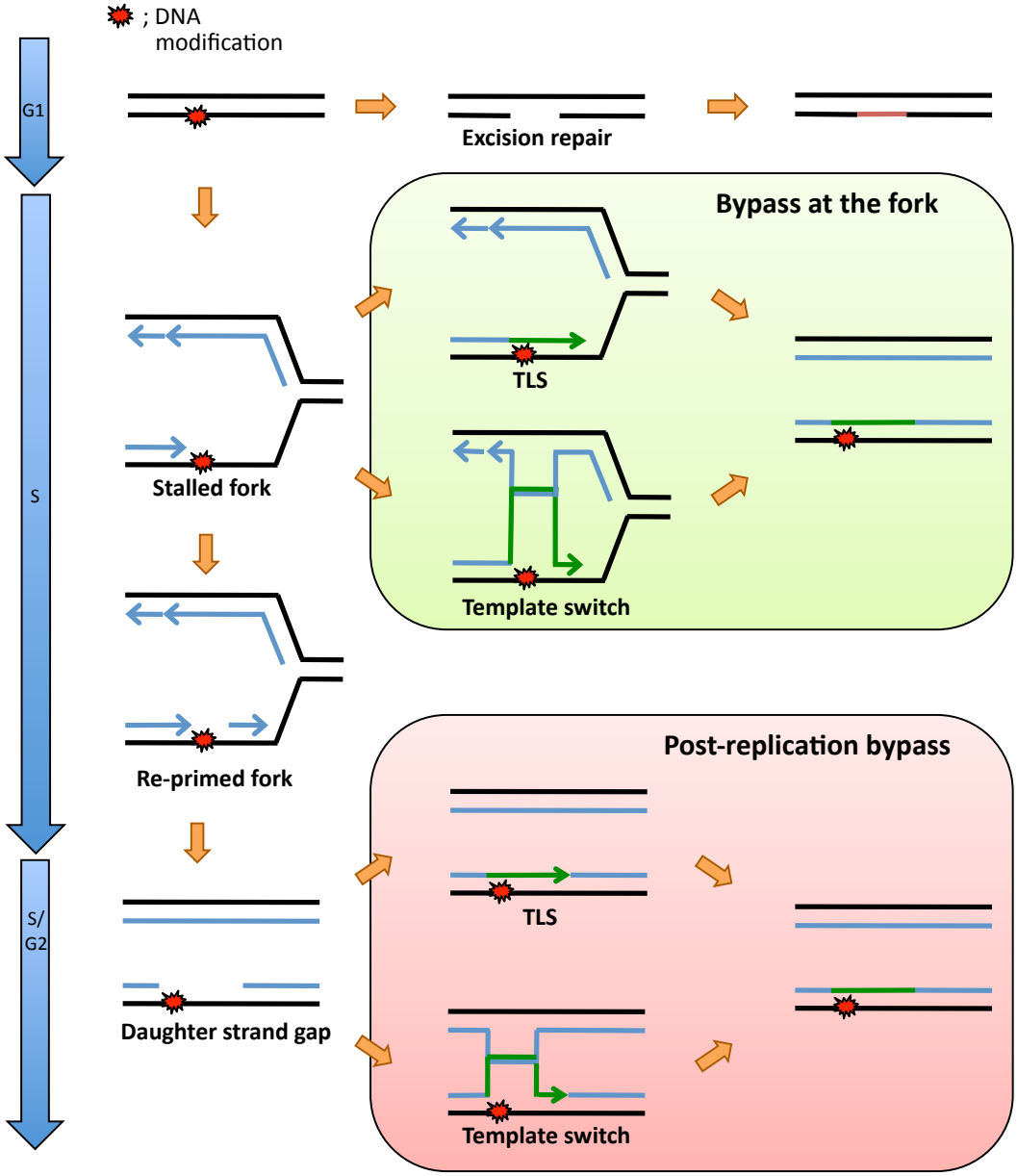


Figure 1.10. Overview of DNA damage tolerance mechanisms in eukaryotic cells.
DNA base lesions are repaired by excision repair mechanisms that use the undamaged strand as a template to restore genetic integrity. It is inevitable that some DNA lesions (depicted on the leading strand above) will escape repair and encounter the DNA replication fork. If this lesion poses a block to the replicative DNA polymerase it will stall the replication fork, and as the DNA lesion is now present in single-stranded DNA, it can no longer be removed by excision repair. Damage tolerance mechanisms (depicted as a green line) exist to overcome this problem, either by using flexible DNA polymerases capable of using the damaged template (trans-lesion synthesis, TLS) or using the lagging-strand as a DNA template (template switch) to overcome the blocked primer terminus. These mechanisms can occur directly at the fork or post-replicatively. In the case of the latter, blockage of the leading strand polymerase causes uncoupling of leading and lagging strand synthesis producing long tracts of single-stranded DNA. DNA synthesis can re-initiate downstream of the lesion on this single-stranded DNA by re-priming, leaving a single-stranded gap encompassing the lesion in the daughter strand. Either TLS or template switching can then be used to fill this daughter strand gap and restore the lesion within duplex DNA, where it can be safely repaired. Note that a lesion on the lagging-strand would result in a daughter strand gap due to discontinuous Okazaki fragment synthesis.

sub-pathways: an error-free pathway including Rad5 and an 'error-prone' pathway with Rad30 and REV3. A number of the genes of the *RAD6* epitasis group were found to be involved in protein ubiquitylation, such as Rad6, Rad18, and Rad5 (Jentsch *et al*., 1987; Bailly *et al*., 1997; Ulrich and Jentsch, 2000), and components of the replication machinery had been implicated in this pathway also, such as PCNA (Torres-Ramos *et al*., 1996). These studies were unified with the key finding that the target of the ubiquitylation proteins was PCNA (Hoege *et al*., 2002). It also became apparent that the error-free pathway of the *RAD6* epistasis group was template switching, whilst the 'error-prone' pathway was TLS. Rev3 and Rev7 were found to be the B-family Pol ζ (Nelson *et al*., 1996a), and Rev1 was discovered to be a deoxycytidine monophosphate (dCMP) transferase (Nelson *et al*., 1996b). Rad30 was also a DNA polymerase, Pol η, and could accurately and efficiently bypass CPDs (Johnson *et al*., 1999a), and was found to be the gene mutated in XP-variant (XP-V) patients (Masutani *et al*., 1999b; Johnson *et al*., 1999b). Further novel DNA polymerases were discovered in all branches of life and together these were classified as the Y-family polymerases (Ohmori *et al*., 2001).

1.4.3. Y-family DNA polymerases and Pol ζ

TLS is a universal mechanism by which cells tolerate DNA damage. It involves the direct replication of damaged DNA by specialised DNA polymerases thus allowing the completion of genome duplication despite the lack of a pristine DNA template. This can occur by the action of a single TLS polymerase, or the combined action of two, with one polymerase inserting opposite the lesion and second extending from the (mis)paired terminus, before the replicative polymerases resume (Figure 1.11). The most abundant group of TLS polymerases belong to the Y-family, which in higher eukaryotes consists of Pol η, ι, κ, and REV1. The B-family Pol ζ is also an important player in TLS. Additionally, various other specialised polymerases from the X-family and a number of novel A-family polymerases have been implicated in this process, but it is the Y-family polymerases that are uniquely adapted for TLS.

1.4.3.1. Features of Y-family DNA polymerases

The key feature of the Y-family DNA polymerases is their ability to replicate past damaged bases, and this is accompanied with low-processivity DNA synthesis and a reduced fidelity on undamaged DNA. Additionally, each Y-family polymerase, particularly those in eukaryotes, appears to be specialised to bypass particular DNA lesions. The first crystallographic structures solved of Y-family polymerases highlighted the features that enable these enzymes to catalyse TLS; they were of Pol η from *S. cerevisiae* (Trincao *et al*., 2001) shortly followed by the archaeal DinB homologue Dpo4, the latter in a ternary

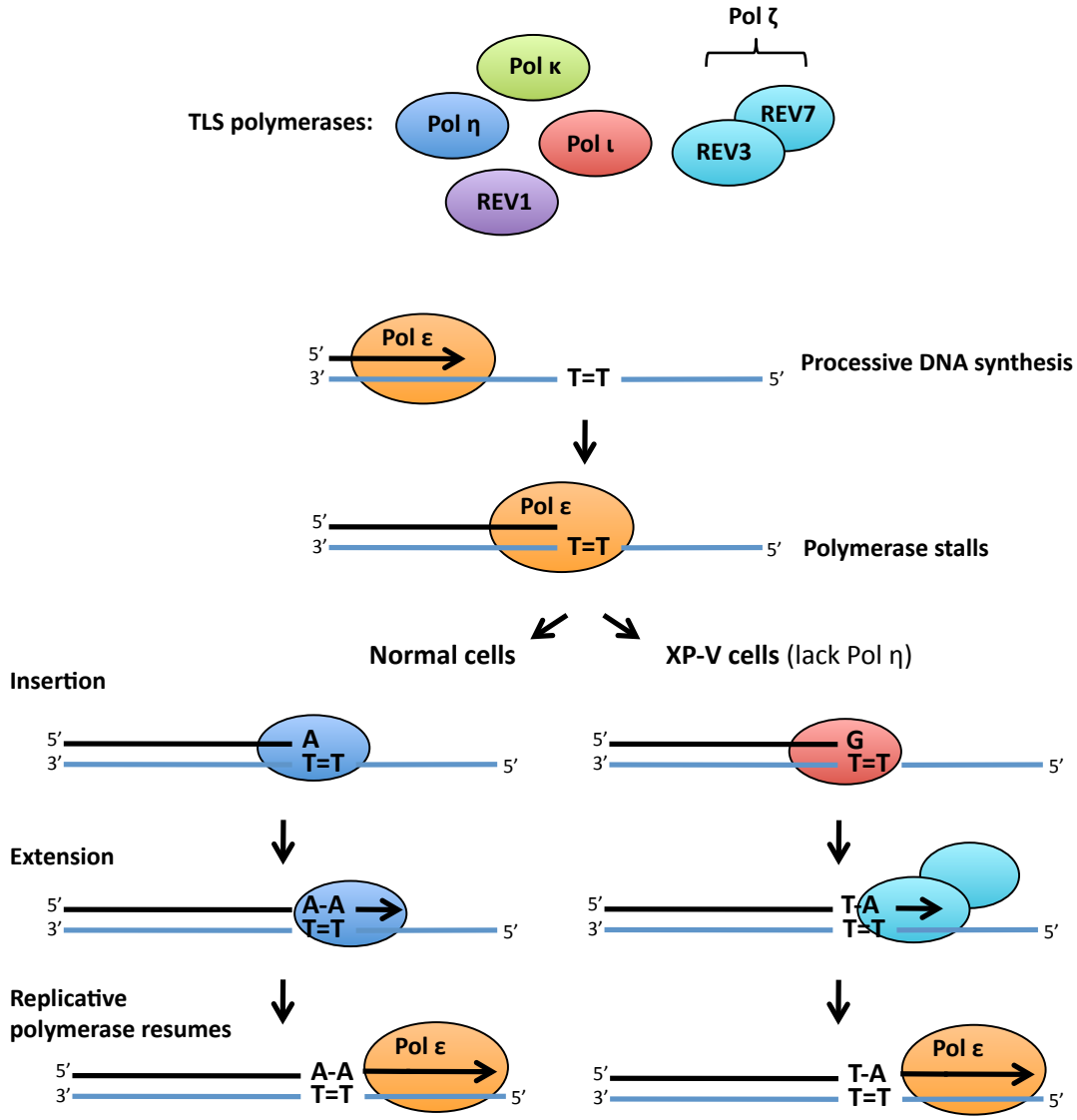


Figure 1.11. One and two polymerase mechanisms of trans-lesion synthesis.
Trans-lesion synthesis (TLS) can occur by both one and two polymerase mechanisms, exemplified above by the bypass of a thymine-thymine cyclobutane pyrimidine dimer (indicated by T=T) in both normal human cells and xeroderma pigmentosum variant (XP-V) patient cells, which lack functional Pol η. Processive DNA synthesis by the replicative polymerase Pol ε is stalled by the presence of a templated cyclobutane thymine dimer. To continue DNA synthesis a polymerase switch occurs, which replaces Pol ε for a TLS polymerase. In normal cells Pol η can efficiently bypass the cyclobutane thymine dimer and correctly inserts and extends from the DNA lesion, to be replaced by Pol ε for the resumption of processive DNA synthesis. In XP-V cells a two polymerase mechanism must be used, with Pol ι (depicted above) or Pol κ mis-incorporating opposite the 3' thymine of the dimer, and Pol ζ incorporating opposite the 5' thymine and then extending before Pol ε resumes. Adapted from Ziv *et al.*, 2009.

complex with DNA and an incoming dNTP (Ling *et al*., 2001). Since then, all the Y-family polymerases in human cells have been co-crystallised with DNA: Pol ι (Nair *et al*., 2004), REV1 (Nair *et al*., 2005), Pol kappa (Lone *et al*., 2007), and Pol η during bypass of a CPD (Biertümpfel *et al*., 2010). As well as these structures documenting the unifying features of the Y-family polymerases, they also display the distinct (and sometimes unusual) ways in which each enzyme is adapted for TLS.

There are several unifying features of the Y-family DNA polymerases. Despite possessing little sequence homology to other DNA polymerases, the catalytic domain of a Y-family polymerase still adopts the right hand topology. The fingers and thumb sub-domains, although of different secondary structure, have the same location in the tertiary structure and their roles in interacting with the incoming dNTP and DNA are conserved, and the palm sub-domain still contains the catalytic carboxylate residues. However, unlike the replicative polymerases of the B- and A-families, the active site cleft of the Y-family polymerases is much more spacious, thus allowing accommodation of bulky adducts on the templating base. Also in contrast to the replicative polymerases are the shorter and stubbier fingers and thumb sub-domains, making fewer contacts with the DNA and incoming nucleotide, which facilitates TLS but contributes to low fidelity and decreased processivity. In addition the Y-family polymerases lack a 3'-5' exonuclease proofreading domain, which will further contribute to decreased fidelity. The Y-family polymerases also posses an additional sub-domain, referred to as a polymerase-associated domain (PAD) (Trincao *et al*., 2001) or "little finger" (Ling *et al*., 2001), the latter in keeping with the human right hand nomenclature. Whilst in the replicative polymerases it is the finger sub-domain that is most mobile (closing with each nucleotide addition) in the Y-family polymerases it is the little finger sub-domain, which together with the thumb sub-domain grasps the DNA. The little finger sub-domain is the least conserved part among the Y-family polymerases catalytic domain, and has been implicated in the lesion specifity of each polymerase. This was demonstrated by swapping the little finger domains of two archeal DinB homologues (Dpo4 and Dbh) and observing that the enzymatic properties of the resulting chimeras were also swapped (Boudsocq *et al*., 2004).

1.4.3.2. DNA polymerase η

Pol η (also called Rad30A, XP-V) is perhaps the best-characterised polymerase of the Y-family and has served as a model for the other members of this family. This is largely due to the clinical relevance of understanding Pol η, as mutations in the gene encoding this polymerase result in the disorder XP-V (Masutani *et al*., 1999b; Johnson *et al*., 1999b).

XP-V is an autosomal recessive disorder characterised by hypersensitivity of the skin to sunlight along with an increased risk of skin cancers (reviewed in Lehmann *et al*., 2011). Whilst the majority of XP disorders result from mutations in any of the seven key NER genes (XP-A to –G), required to repair UV light induced DNA lesions, XP-V results from an inability to efficiently replicate UV damaged DNA (Lehmann *et al*., 1975). This is consistent with the ability of Pol η to accurately bypass *cis-syn* cyclobutane thymine dimers (Johnson *et al*., 1999; Masutani *et al*., 1999a), the most common lesion generated by UV in DNA, which it does with a similar efficiency to replicating undamaged DNA (McCulloch *et al*., 2004). This suggests Pol η is largely responsible for the bypass of these lesions *in vivo*. The greater frequency of UV induced mutations and altered UV mutations spectrum of XP-V patients will result from other, less accurate, DNA polymerases performing TLS of CPDs. This has been suggested to be Pol ι (Dumstorf *et al*., 2006) along with Pol κ and ζ (Ziv *et al*., 2009) (Figure 1.11). The extent of Pol η's specialisation to replicate past CPDs was recently demonstrated in a set of crystal structures (Silverstein *et al*., 2010; Biertümpfel *et al*., 2010). The active site of Pol η is large, capable of accommodating both thymines of the CPD, and the linked thymines are stabilised to allow correct pairing with adenine. As the CPD remains in duplex DNA, problems can arise following TLS of the dimer, as the CPD distorts the DNA making it possible for slippage of the polymerase to occur producing frameshift mutations. However, the little finger domain of Pol η acts as a molecular splint to stabilise the CPD containing DNA into normal B-form conformation. Additionally, to prevent the low-fidelity of Pol η (Matsuda *et al*., 2000) introducing mutations beyond the lesion, following replication of three bases, steric clashes occur with the DNA ensuring the dissociation of Pol η (Biertümpfel *et al*., 2010).

Although Pol η is unable to bypass the other major UV DNA lesion, the 6-4 photoproduct, stalling following incorporation of a single base opposite the 3' T (Masutani *et al*., 2000; Johnson *et al*., 2001), Pol η can catalyse TLS of a number of other lesion *in vitro* (reviewed in Lehmann *et al*., 2002), although less efficiently than a CPD. For example Pol η can efficiently bypass an 8-oxo-G *in vitro* (Haracska *et al*., 2000; Zhang *et al*., 2000c), although this does not appear to be an *in vivo* function of this enzyme (Avkin and Levneh, 2002). Pol η can also catalyse TLS across a cisplatin adduct (Masutani *et al*., 2000) and XP-V cells are sensitive to cisplatin suggesting this could be an additional role of Pol η (Albertella *et al*., 2005). In addition to TLS of DNA lesions, Pol η has also been implicated in homologous recombination (McIlwraith *et al*., 2005), DNA replication during an unperturbed S-phase, possibly in the replication of chromosomal fragile sites

(Rey *et al*., 2009), gene hypermutation during development of the immune system (Zeng *et al*., 2001), and the repair of clustered oxidative damage (Zlatanou *et al*., 2012).

1.4.3.3. DNA polymerase ι

Pol ι (also called Rad30B) was identified as a paralogue of Pol η (McDonald *et al*., 1999), and is a highly distributive and low-fidelity polymerase (Tissier *et al*., 2000a). The extent of Pol ι's participation in TLS of UV photoproducts is a debated topic (reviewed in Vidal and Woodgate, 2009); with regards to the CPD lesion some studies have reported that Pol ι is completely blocked (Johnson *et al*., 2000b; Zhang *et al*., 2000a), whilst other studies have reported incorporation opposite the 3' T followed by limited complete bypass (Tissier *et al*., 2000b). With regards to the 6-4 photoproduct, Pol ι has been suggested to be involved in a two-step mechanism, by which it incorporates opposite the lesion and a second TLS polymerase, namely Pol ζ, extends from the mismatched terminus (Tissier *et al*., 2000b; Johnson *et al*., 2000b). Various studies have demonstrated that Pol ι is required for the bypass of a subset of UV-induced DNA lesions *in vivo* (Ohkumo *et al*., 2006; Dumstorf *et al*., 2006; Gueranger *et al*., 2008) with its role becoming more apparent with the lack of Pol η in XP-V cells (Ziv *et al*., 2009).

Pol ι has a particularly intriguing feature in that it exhibits over a thousand-fold difference in fidelity depending on the templating base, with a modest fidelity opposite a templated adenine but a high error rate opposite a templated thymine (Zhang *et al*., 2000a). The reason for this is due to residues in the finger domain of Pol ι that restricts the position of the templating base preventing a Watson-Crick pairing, instead promoting Hoogsteen base pairing, which uses alternative hydrogen bond donors/acceptors on the purine rings altering the geometry of the double helix (Nair *et al*., 2004). This feature enables Pol ι to correctly incorporate opposite the important oxidative lesion 8-oxo-G, as it restricts the geometry of the 8-oxo-G promoting the formation of the most stable and correct pair with dC (Kirouac *et al*., 2011). Pol ι has also been implicated in repair of oxidative lesions through the BER pathway. It was discovered that Pol ι has 5'-deoxyribose phosphate (dRP)-lyase activity, which coupled with its polymerase activity and the addition of purified BER components, could successfully reconstitute *in vitro* BER of uracil containing DNA (Bebenek *et al*., 2001). Similarly, the activity of Pol ι has been shown to complement BER-deficient cell extracts lacking Pol β (Prasad *et al*., 2003). A more recent study has further demonstrated Pol ι's role in protecting cells against the effects of oxidative damage; showing that Pol ι knockdown cells are hypersensitive to oxidative damage coupled with a reduced BER activity, and that Pol ι

is recruited to sites of oxidative damage and interacts with X-ray repair cross-complementing protein (XRCC)1, a scaffold for BER components. (Petta *et al*., 2008).

1.4.3.4. DNA polymerase κ

Pol κ (also called DinB) was identified by homology searching for eukaryotic homologues of the *E. coli DinB* gene (Ogi *et al*., 1999; Gerlach *et al*., 1999; Johnson *et al*., 2000a), which encodes the *E. coli* TLS polymerase Pol IV, and is the only Y-family polymerase conserved throughout all domains of life (Ohmuri *et al*., 2001). Like all Y-family polymerases it is a low-fidelity DNA polymerase, however it is the most accurate of its family (Johnson *et al*., 2000a). This is thought to be due to an additional feature of the Pol κ catalytic domain, an amino-terminal clasp (called N-clasp), which makes additional interactions with DNA allowing Pol κ to encircle the DNA molecule (Lone *et al*., 2007). In addition to the relatively restricted active site of Pol κ (in comparison to other Y-family members) this partly explains the differing TLS capabilities of this Y-family polymerase (Lone *et al*., 2007). Pol κ is unable to incorporate opposite UV photoproducts and incorporates with a low efficiency opposite a number of other lesions *in vitro* (Zhang *et al*., 2000b; Ohashi *et al*., 2000). However, it has been shown to be an efficient extender of mispaired termini (Washington *et al*., 2002; Haracska *et al*., 2002b), and this extender function of Pol κ is particularly apparent in XP-V cells (Ziv *et al*., 2009). A group of DNA lesions that Pol κ can efficiently bypass are those caused by polycyclic aromatic hydrocarbons, which are ubiquitous environmental carcinogens and are present in tobacco smoke (Phillips, 1983). One such example is benzo[*a*]pyrene-guanine, with bypass being demonstrated *in vitro* (Zhang *et al*., 2000b) and *in vivo*, confirming this is a cellular role of Pol κ (Ogi *et al*, 2002; Avkin *et al*., 2004). Pol κ has also been implicated in replicating structured DNA *in vivo*, which can also impede DNA replication (Bétous *et al*., 2009)

Despite Pol κ only having a possible minor role in the bypass of UV photoproducts *in vivo*, cells deficient in Pol κ are sensitive to UV light, and it was found that these cells actually had a deficiency in NER (Ogi and Lehmann, 2006). It was later demonstrated that Pol κ is required for the DNA synthetic step during NER, functioning in the same pathway as Pol δ and together accounting for half of NER activity (with the other half being Pol ε) (Ogi *et al*., 2010). It has been speculated that Pol κ is required for NER during times of low dNTP concentrations or when the DNA template contains difficult to replicate structures (Ogi *et al*., 2010; Lehmann, 2011).

1.4.3.5. REV1

REV1 is the only member of the Y-family to escape the Greek letter nomenclature. This is because REV1 is not a DNA polymerase but a dCMP transferase, capable of incorporating dCMP opposite templated guanines (Nelson *et al*., 1996). REV1's TLS capabilities are limited, being able to efficiently incorporate opposite bulky adducted guanines (Zhang *et al*., 2002) and abasic sites (Nelson *et al*., 1996). Bypass of abasic sites has been shown to be an *in vivo* function of REV1, as without the catalytic activity of REV1 the mutation spectrum is altered in vertebrate cells immunoglobulin gene hypermutation (Masuda *et al*., 2009; Ross and Sale, 2006; Jansen *et al*., 2006) and during bypass of endogenous abasic sites in yeast (Mudrak and Jinks-Robertson, 2011). The crystal structure of REV1 revealed how this enzyme is adapted for its cellular role. Rather than using Watson-Crick bonding to detect the correct pairing of the incoming dCMP and the templated dG, residues in the little finger of REV1 temporarily coordinate the templated dG and actually flip this base out, whilst the incoming dCMP hydrogen bonds with an arginine residue. The base flipping creates space to accommodate bulky guanine adducts and the hydrogen bonding to the incoming dCMP remove the need for a guanine base for correct incorporation (Nair *et al*., 2007).

REV1, like the other REV genes, is required for UV-induced mutagenesis in yeast (Lemontt, 1971) and also humans (Gibbs *et al*., 2000). Although curiously, the catalytic activity of REV1 is not required for mutagenesis, as without catalytic activity the mutagenesis still occurs but the mutation spectrum is altered (Otsuka *et al*., 2005). This potential non-catalytic role of REV1 was first noted by Nelson and colleagues in yeast, as REV1 was essential for *in vivo* bypass of a 6-4 photoproduct yet could not insert a single nucleotide opposite the lesion *in vitro* (Nelson *et al*., 1996; Nelson *et al*., 2000). The carboxyl-terminus of REV1 interacts with all other Y-family polymerases and Pol ζ (Murakumo *et al*., 2001; Guo *et al*., 2003; Ohashi *et al*., 2004; Tissier *et al*., 2004) and the amino-terminus contains a BRCT (BRCA1 carboxyl-terminus) domain that is important for PCNA interaction and survival following exposure to UV light (Guo *et al*., 2006a). These studies of REV1 show the alternate role is mediated by protein-protein interaction and that REV1 acts as a binding platform for other TLS polymerases, playing an important regulatory role in TLS (discussed further below). Other reported roles for REV1 include immunoglobulin gene somatic hypermutation as mentioned previously, the replication of G-quadruplex DNA (Sarkies *et al*., 2010), and HR (Sharma *et al*., 2012)

1.4.3.6. DNA polymerase ζ

Pol ζ is a member of the B-family polymerases and was originally characterised as a heterodimer composed of a catalytic (Rev3 in yeast, Rev3L in higher eukaryotes) and accessory (Rev7) subunit (Nelson *et al*., 1996b). However recent studies in yeast have shown that Pol ζ consists of four subunits, sharing accessory subunits Pol31 and Pol32 (p50 and p66 in humans) from the replicative Pol δ. These were demonstrated as essential subunits in yeast Pol ζ (Johnson *et al*., 2012) and another study suggests this also may be the case in human cells (Baranovskly *et al*., 2012). Our understanding of the *in vitro* TLS capabilities of Pol ζ come from experiments with the yeast enzyme, as an active form of mammalian Pol ζ (with REV3L being over 3000 amino acids) remains to be purified. Compared to the Y-family polymerases, the fidelity of DNA synthesis by Pol ζ is relatively high (Zhong *et al*., 2006), and more akin to that of Pol α (Thomas *et al*., 1991). Also like Pol α, Pol ζ does not contain a 3'-5' exonuclease proofreading activity. Pol ζ can bypass a number of DNA lesions completely, such as a CPD (Nelson *et al*., 1996b) and thymine glycol (Johnson *et al*., 2003), but with a low efficiency. Rather, it is the extension of mispaired termini that Pol ζ appears to be specialised for. Pol ζ can extend from bases paired opposite a DNA lesion often with the same efficiency as if the base was paired with undamaged DNA, exemplified by extension from a CPD, 6-4 photoproduct (Johnson *et al*., 2000a) and thymidine glycol (Johnson *et al*., 2003). This extension from lesions reflects an *in vivo* role of Pol ζ in mammalian cells, demonstrated for 6-4 photoproducts (Yoon *et al*., 2010a) and thymidine glycol (Yoon *et al*., 2010b).

One striking difference between Pol ζ and the Y-family polymerases is that disruption of the Pol ζ (*REV3L*) gene is embryonic lethal in mice, demonstrating an essential role in mammalian embryonic development (Wittschieben *et al*., 2000; Esposito *et al*., 2000; Bemark *et al*., 2000). Pol ζ is essential for normal cell proliferation in mammalian cells, with the bypass of endogenous oxidative damage being partly responsible (Lange *et al*., 2012). In yeast, Pol ζ has been demonstrated to be required for the post-replication damage tolerance of genomic rNTPs (Lazzaro *et al*., 2012) that are frequently misincorporated by the replicative DNA polymerases (Nick McElhinny *et al*., 2010). This is known to be a plentiful source of endogenous DNA damage in mammalian cells (Reijns *et al*., 2012) and may also contribute to the essential function of Pol ζ. In addition to its roles in TLS, Pol ζ has been shown to participate in the repair of double-strand breaks, along with REV1, via HR (Sharma *et al*., 2012).

1.4.4. The PCNA switchboard and the regulation of TLS

The Y-family polymerases η, ι, κ, and REV1, are all nuclear proteins and localise to sites of active DNA replication during S-phase, called replication factories (Kannouche *et al*., 2001, 2003; Tissier *et al*., 2004; Ogi *et al*., 2005). Although they are in close proximity to the replication fork their access to the primer terminus must be tightly regulated to ensure these low-fidelity polymerases are only used when required; and the post-translational modification of the sliding clamp PCNA is central to this regulation (Figure 1.12). Upon encountering a bulky DNA adduct (on the leading strand for example) the replicative DNA polymerase will stall due to an inability to catalyse DNA synthesis and the MCM helicase will continue to unwind duplex DNA. This functional uncoupling of the replicative helicase and polymerase produces long stretches of unwound primed single-stranded DNA that will become coated with RPA (Byun *et al*, 2005). RPA-coated primed single-stranded DNA is the trigger for ubiquitination of PCNA (Change *et al*., 2006). RPA recruits and interacts with the ubiquitin-ligase (E3) Rad18, which together with its cognate ubiquitin-conjugator (E2) Rad6, mono-ubiquitinate PCNA on the conserved lysine-164 residue (Davies *et al*., 2008). It is this mono-ubiquitination that activates the so-called 'error-prone' TLS pathway (Stelter and Ulrich, 2003). Additionally, a second E2-E3 pair, the E2 Ubc12-Mms2 and the E3 Rad5 (in yeast), use the mono-ubiquitination as a substrate for poly-ubiquitination via the lysine-63 linkage of ubiquitin (Davies *et al*., 2008). In *S. cerevisiae* this activates the error-free template switching pathway (Stelter and Ulrich, 2003), however in mammalian cells it is more complex and so less well understood. Although evidence of non-TLS damage tolerance, mediated by poly-ubiquitinated PCNA (suggestive of template switching), does exist in mammalian cells (Chui *et al*., 2006), following treatment of cells with replication fork blocking agents it is mono-ubiquitantion of PCNA that is predominant (Kannouche *et al*., 2004; Chui *et al*., 2006). Further, mammalian homologues of Rad5, SHPRH and HLTF, are both involved in the TLS pathway (Lin *et al*., 2011). Together these reports suggest it is TLS that is the dominant mechanism of DNA damage tolerance in human cells.

A number of interactions between the Y-family DNA polymerases and PCNA mediate the switch from replicative to TLS polymerase. Pol η, ι, and κ, contain PCNA-interacting protein (PIP) motifs at their extreme carboxyl-terminus that allow interaction with PCNA (Haracska *et al*., 2001a, 2001b, 2002a), and these have been shown to be partly responsible for the localisation of these polymerases at replication forks stalled at DNA lesions (Kannouche *et al*., 2001, 2003; Tissier *et al*., 2004; Vidal *et al*., 2004; Ogi *et al*., 2005). Additionally, the mono-ubiquitination of PCNA upon replication fork stalling increases affinity of the sliding clamp for the Y-family polymerases, which was

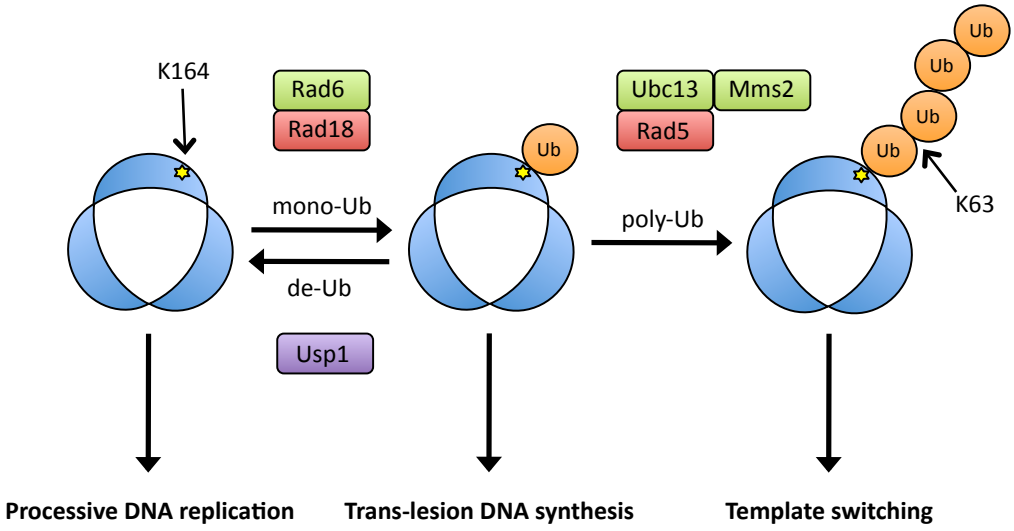


Figure 1.12. Regulation of DNA damage tolerance by the DNA sliding clamp PCNA.
The post-translational modification of the polymerase sliding clamp PCNA by the covalent attachment of ubiquitin (Ub, orange circle), is key to the regulation of DNA damage tolerance in eukaryotic cells (*S. cerevisiae* shown). The enzymes responsible are depicted above: the E2 ubiquitin conjugator (green box), E3 ubiquitin ligase (red box), and de-ubiquitination enzyme (DUB) (purple box). Stalling of the replication fork produces primed RPA-coated single-stranded DNA that recruits the E3 Rad18 and E2 Rad6, which mono-ubiquitinate the PCNA homotrimer on a conserved lysine 164 residue. This leads to the recruitment of trans-lesion synthesis DNA polymerases that catalyse both error-free and mutagenic damage bypass. Mono-ubiquitinated PCNA is also the substrate for a lysine 63-linked poly-ubiquitin chain by the concerted actions of the E2 Ubc13-Mms2 and E3 Rad5. This promotes error-free damage tolerance by template switching. In the absence of DNA lesions, the DUB USP1 removes ubiquitin from mono-ubiquitinated PCNA thereby promoting processive DNA synthesis. Although only one subunit of the PCNA homotrimer is shown to be modified, it is likely that all three and ubiquitinated *in vivo*.

demonstrated initially for Pol η (Kannouche *et al*., 2004; Watanabe *et al*., 2004). All Y-family polymerases contain ubiquitin-binding motifs (UBM), which in the case of Pol η and κ are zinc-fingers and so are called ubiquitin-binding zinc-fingers (UBZs). These promote the interaction of the Y-family polymerases with mono-ubiquitinated PCNA and are therefore critical to their role in DNA damage tolerance (Bienko *et al*., 2005; Plosky *et al*., 2006; Guo *et al*., 2006b, 2008). In addition to facilitating access of the TLS polymerase to the primer terminus, mono-ubiquitinated PCNA can also specifically increase the efficiency of these polymerases on damaged DNA. *In vitro* experiments have demonstrated that whilst mono-ubiquitination of PCNA does not change its properties as a processivity factor for Pol δ, ε, and even Pol η on undamaged DNA, the TLS efficiency of Pol η and REV1 were substantially increased on an abasic site containing template (Garg and Burgers, 2005b). Selection of the required TLS polymerase at the primer terminus remains an important question. However given the different bypass capabilities of the TLS polymerases it may be simply down to enzyme kinetics, i.e. whether or not the polymerase can catalyse DNA synthesis and then dissociate from the template. Responsible for removing ubiquitin from PCNA is the de-ubiquitination enzyme (DUB) ubiquitin specific protease 1 (USP1) (Huang *et al*., 2006). DUBs are cysteine proteases that specifically cleave ubiquitin conjugates (Amerik and Hochstrasser, 2004). USP1 keeps the levels of ubiquitinated PCNA low in undamaged cells and undergoes an auto-cleavage reaction following treatment of cells with UV radiation (Huang *et al*., 2006).

This simple and elegant model of TLS regulation in eukaryotes, summarised in Figure 1.12, is not the complete picture however, as further investigations have revealed complexities. For instance many different proteins affect the ubiquitination of PCNA, such as the checkpoint kinase Chk1 (Yang *et al*., 2008), p21 and p53 (Avkin *et al*., 2006; Soria *et al*., 2006), and chromatin remodelers such as BAF180 (Niimi *et al*., 2012). Also, a recent study has shown that a subset of TLS in mammalian cells occurs independently of PCNA ubiquitination, and is instead partly mediated by protein-protein interactions with the Y-family polymerase REV1 (Hendel *et al*., 2011). Additionally, there are conflicting reports on the necessity of the ubiquitin-binding motifs of TLS polymerases. Reports have shown that the UBZ of Pol η is required for correct localisation following UV irradiation, complementation of the XP-V phenotype, and *in vitro* bypass of a CPD using cell extracts (Bienko *et al*., 2005; Plosky *et al*., 2006; Sabbioneda *et al*., 2009; Schmutz *et al*., 2010). However a number of studies from the Prakash laboratory suggest the UBZ of Pol η is dispensable for this enzymes function in both yeast and human cells. A recent study has attempted to clarify these discrepancies and concluded that both the PIP box and UBZ motif co-operate to retain Pol η at stalled replication forks, and whilst individual

mutation of each domain results in slight phenotypes, mutation of both renders cells with XP-V characteristics (Despras *et al*., 2012). In addition to the post-translational modifcation of PCNA playing an important regulatory role in TLS, the Y-family polymerases themselves are also targets for modification. Pol η is mono-ubiquitinated and this is suggested to promote intermolecular interactions with its UBZ (Bienko *et al*., 2010) and also the UBM of Pol ι (McIntyre *et al*., 2012), thereby preventing unwanted interaction with PCNA. Pol η is also phosphorylated by the intra-S checkpoint kinase ATR, which has been demonstrated to be important for its crucial role in the bypass of UV photoproducts (Göhler *et al*., 2011).

1.4.5. Fork-associated or post-replicative DNA damage tolerance?

The timing of lesion bypass has been a long-standing question in the TLS field. More precisely, whether TLS is associated with chromosomal DNA replication or is separate, occurring after the replication fork has progressed (Figure 1.10) (reviewed in Lehmann and Fuchs, 2006). Early experiments suggested bypass occurred post-replicatively (Rupp and Howard-Flanders, 1968; Lehmann 1972) whilst since then models have focussed on bypass *in situ* at blocked forks. These models largely arose due to the semi-discontinuous nature of DNA synthesis. A lesion on the lagging strand would be predicted not to block progression of the replication fork due to the discontinuous synthesis of this strand, as the re-initiation of a downstream Okazaki fragment would allow progression and a gap would be left behind. A lesion on the leading strand however, would be predicted to inhibit DNA synthesis to a greater extent owing to the continuous synthesis of this strand. This differing leading and lagging strand synthesis of damaged DNA has been observed *in vitro* (Svoboda and Vos, 1995). However, the initial report of DNA damage tolerance in *E. coli* by Rupp and Howard-Flanders clearly demonstrated that a DNA lesion does not block bulk DNA synthesis (Rupp and Howard-Flanders, 1968). This would be consistent with discontinuous synthesis on the leading strand also. Upon the replication fork encountering a leading strand lesion and following uncoupling of the replicative helicase and polymerase, DNA synthesis would be re-initiated downstream of the lesion leaving a single-stranded gap in the daughter strand. The mechanism of replication restart by re-priming has been elegantly demonstrated in bacteria *in vitro* (Heller and Marians, 2006; Yeeles and Marians, 2012), however this has not been clearly demonstrated in eukaryotes. A number of studies imply this process occurs (Lehmann, 1972; Jansen *et al*., 2009b; Elvers *et al*., 2011), and perhaps most convincingly, the single-stranded gaps produced by re-priming of damaged DNA have been visualised *in vivo* by electron microscopy in yeast (Lopez *et al*., 2006). Further studies in yeast have shown that DNA damage tolerance can occur separately from

chromosomal replication with no adverse effects (Daigaku *et al*., 2010; Karras *et al*., 2010), whilst TLS in mammalian cells has been shown to occur in G2 phase as well as during S-phase, interestingly with a different mutational spectrum (Diamant *et al*., 2012).

Fork-associated bypass and post-replicative bypass were first genetically distinguished in chicken cells. Mono-ubiquitination of PCNA was required for post-replicative bypass and bypass at the fork required the Y-family polymerase REV1, specifically the carboxyl-terminus of REV1 that is known to interact with other Y-family polymerases (Edmunds *et al*., 2008). This genetic distinction, however, has not held true for mammalian cells. Although mono-ubiquitinated PCNA is involved in post-replicative bypass (Niimi *et al*., 2008), and evidence of an 'early' (fork-associated) and 'late' (post-replicatve) TLS pathway has been shown in mammalian cells, REV1 and ubiquitination of PCNA were involved in both pathways (Jansen *et al*., 2006; Temviriyanukul *et al*., 2012). This shows that there are clearly differences in the regulation of TLS between eukaryotes. It is likely that the timing of damage bypass is dependant on a number of variables, including the location of the lesion (leading versus lagging) and the nature of the lesion (reviewed in Sale, 2012; Daigaku, 2012). It can be envisaged that if the DNA lesion on the leading strand is efficiently bypassed by TLS then perhaps it is bypassed quickly and directly at the fork, allowing continued progression of this fork. However if the lesion is bypassed inefficiently and therefore takes longer, the uncoupled helicase will progress further producing more single-stranded DNA on which re-priming could occur. This will produce a gap encompassing the lesion on the daughter strand that will be susbsequently filled by damage tolerance mechanisms behind the fork (Figure 1.10).

1.5. The African Trypanosome - *Trypanosoma brucei*

Trypanosoma brucei, the African trypanosome, along with the closely related *Trypanosoma cruzi* and the *Leishmania*'s, are protozoan parasites of significant medical importance. *T. brucei* is the causative agent of human African trypanosomasis, also called African sleeping-sickness, a typically fatal condition endemic in sub-Saharan Africa, whilst *T. cruzi* and the *Leishmania*'s and responsible for Chagas disease (American trypanosomasis) and Leshmania respectively. Together these pathogens affect more than 20 million people worldwide and cause approximately 110,000 deaths annually (World Health Organisation, 2008). These three unicellular microbes are members of the family Trypanosomatidae of the order Kinetoplastida, and are characterised by a single flagellum and a mitochondrion whose genome is arranged in a unique and complex catanated network termed a kinetoplast. In addition to their medical importance, trypanosomatids offer a unique evolutionary perspective as they are

amongst the earliest diverging organisms from the eukaryotic tree (Simpson *et al*., 2006). As a result, these morphologically rather simple organisms possess a number of features considered peculiar among eukaryotes, including polycistronic transcription of most their genome and *trans*-splicing of all mRNA transcripts (Martínez-Calvillo *et al*., 2010), the previously mentioned unique mitochondrial DNA architecture along with a complex method of mitochondrial RNA editing (Liu *et al*., 2005; Lukeš *et al*., 2005), and the various mechanisms used by these parasites to escape their host's immune system (Machado *et al*., 2006; Horn and McCulloch, 2010).

1.5.1. African sleeping sickness

T. brucei is an extracellular parasite and is spread among its mammalian hosts by the blood-feeding tsetse flies of the genus *Glossina*. The tsetse fly is restricted to the African continent from south of the Sahara to north of the Kalahari Desert, and thus confines the transmission of *T. brucei* and the distribution of the disease. Three morphologically identical sub-species of *T. brucei* exist; two infect humans, *T. brucei gambiense* and *rhodesiense*, and a third, *T. brucei brucei*, infects cattle causing Nagana (animal trypanosomaisis). *T. brucei gambiense* is largely in central and west Africa and results in a slow and chronic disease, whilst *T. brucei rhodesiense* is largely in east and southern Africa and results in a more acute disease. If untreated, human African trypanosomaisis in invariably fatal. The disease progresses in two distinct stages; first is the haemolymphatic phase in which the parasites invade the bloodstream and lymphatic system which results in fever, headaches, and swelling of the lymph nodes. Second is the neurological phase in which the parasites cross the blood-brain barrier and invade the central nervous system. This disrupts circadian rhythms leading to altered sleep patterns (giving rise to the name sleeping sickness) and eventually leads to coma and death (reviewed in Barrett *et al*., 2003; Brun *et al*., 2010).

To facilitate the cyclical transmission between its mammalian hosts and fly vector, *T. brucei* employs a complex life cycle with a number of distinct phases involving morphological, surface coat, and biochemical changes (reviewed in Matthews *et al*., 2004). When in the tsetse salivary gland the trypanosomes are in their infectious non-dividing metacyclic form, and when the infected tsetse fly bites, the trypanosomes are inoculated into the mammalian host. These trypanosomes then re-enter the cell cycle and become the morphologically long-slender form and begin to express the bloodstream-stage specific antigens, called variant surface glycoproteins (VSGs), which allow these parasites to evade the host's immune system. When the parasite density becomes high the trypanosomes stop dividing and become the morphologically stumpy

form, which is pre-adapted for the tsetse fly environment but has a limited life span in the host's bloodstream. This is thought to prevent the trypanosome infection killing the host and therefore increasing the probability of transmission. When a tsetse fly bites the infected host, the stumpy-form trypanosomes are ingested in the bloodmeal into the gut and again re-enter the cell-cycle and differentiate into the procyclic form, shedding their VSG coat for a procyclic surface antigen coat. The parasites then migrate to the salivary gland of the tsetse fly and eventually mature to the non-dividing metacyclic form. In the laboratory the non-human infective *T. brucei brucei* is commonly used in either its procyclic form or the pathogenic long-slender bloodstream-form, with the latter being exclusively used in this thesis.

1.5.2. Organisation and replication of the nuclear genome

In contrast to higher eukaryotes, our understanding of nuclear DNA metabolism in trypanosomes pales in comparison to its mitochondrial counterpart, particularly with respect to DNA replication. In a similar vain to the mitochondrial genome, the nuclear genome of trypanosomatids has a novel organisation. This was first observed using pulse-field gel electrophoresis (Van der Ploeg *et al*., 1984a and 1984b), as unlike other eukaryotes, the nuclear DNA of trypanosomatids does not condense during mitosis and thus prevents determination of the chromosomal karyotype by cytological methods. The nuclear chromosomes of *T. brucei* can be divided into three different classes based on their migration during pulse-field gel electrophoresis: the megabase chromosomes (1-6 Mb), the intermediate chromosomes (200-900 kb), and the minichromosomes (30-150 kb). The megabase chromosomes are diploid, with 11 pairs, and contain the majority of the protein-coding genes, whilst there are approximately 5 intermediate chromosomes and 100 minichromosomes, which are probably anueploid (El-Sayed *et al*., 2000; Daniels *et al*., 2010).

The replication fork machinery responsible for duplicating the nuclear genome is largely thought to resemble the higher eukaryotic machinery, given that much of the components are conserved in trypanosomatid genomes (El-Sayed *et al*., 2005). However, there are a number of differences, particularly in the initiation of DNA replication. Analysis of the whole genome sequences of *T. brucei* (Berriman *et al*., 2005), *T. cruzi* (El-Sayed *et al*., 2005), and *L. major* (Ivens *et al*., 2005), collectively called the TriTryp's, identified only a single ORC subunit rather than the six present in higher eukaryotic genomes, and additionally the ORC subunit had homology to CDC6, which is higher eukaryotes is a distinct protein. This led to the suggestion that initiation of DNA replication in trypanosomes could resemble the archaeal mechanism (El-Sayed *et al*., 2005), as

archaeal genomes encode a single ORC1-CDC6 protein (Kelman and Kelman, 2003). This was further suggested following initial characterisation of the *T. brucei* and *T. cruzi* ORC1-CDC6 protein (Godoy *et al*., 2009). However, subsequent studies have identified additional putative divergent ORC homologues, some of which are so divergent they could represent trypanosomatid innovations. It remains to be established whether the five potential ORC homologues form a complex as in higher eukaryotes but they do interact with each other, and knockdown of these proteins results in similar phenotypes that could be consistent with a role in DNA replication (Dang and Li, 2009; Tiengwe *et al*., 2012a). Despite the divergence of the ORC proteins, the TriTryp genomes encode the six subunits of the MCM2-7 helicase complex (El-Sayed *et al*., 2005), which was demonstrated to form a heterohexamer as in higher eukaryotes (Tiengwe *et al*., 2012a). Similarly the MCM helicase complex interacts with some of the ORC homologues (Dang and Li, 2009; Tiengwe *et al*., 2012a) and forms an active DNA helicase when complexed with the trypanosome CDC45 and GINS complex (Dang and Li, 2009), as in higher eukaryotes. This suggests the divergence of trypanosomatid replication machinery may be restricted to the early steps of DNA replication initiation. It is notable that although the three replicative B-family DNA polymerases, α, δ, and ε, are present in the TriTryp genomes, homologues of the some of the accessory subunits for each of these polymerases were not identified (El-Sayed *et al*., 2005), suggesting these proteins have diverged significantly or perhaps the replicative polymerases in trypanosomatids exist as different complexes. The catalytic subunits of each these polymerases were identified as essential genes in a recent genome-wide RNAi screen (Alsford *et al*., 2010), which would be consistent with a role in DNA replication.

A recent study significantly advanced our current understanding of DNA replication in *T. brucei*, by mapping the replication origins in the 11 diploid megabase chromosomes through determining ORC1-CDC6 binding sites (Tiengwe *et al*. 2012b). This study revealed a striking co-ordination of DNA replication and transcription that is unprecedented in eukaryotes. The origins of replication were located to the boundaries of the transcribed domains, and further, ORC1-CDC6 binding to these locations helped define these transcription boundaries. The investigators speculated that this may be due to the novel arrangement of the protein-coding genes in *T brucei*, which are arranged in clusters that on average contain 50 or more genes that are constitutively transcribed in a polycistronic fashion, and subsequently regulated on a post-transcriptional level (reviewed in Daniels *et al*., 2010). Another striking observation from this study was the scarcity of replication origins required to completely replicate the 11 megabase chromosome pairs, being predicted as fewer than 100. It was also noted that, unlike

bacteria and other eukaryotes, there was several instances were DNA replcation and transcription were opposed, which is known to be a cause of genetic instability (reviewed in Bermejo *et al*., 2010). It is clear from these and other studies that although our understanding of DNA replication is advancing in these important human pathogens, there is much further work to be done.

1.5.3. DNA repair and damage tolerance

DNA repair is of particular relevance to trypanosomatids, given the hostile environment of their mammalian hosts in which these parasites proliferate. Additionally, both *T. brucei* and *T. cruzi* have co-opted DNA repair mechanisms, HR and MMR respectively, to create genetic diversity that is used to escape from their host's immune system. The TriTryp genomes encode many key DNA repair proteins suggesting these parasites are capable of catalysing most repair mechanisms (El-Sayed *et al*., 2005), however there are differences. The trypanosomatid genomes encode enzymes capable of directly repairing oxidative and alkylation DNA damage, with the recent preliminary characterisation of a *T. brucei* AlkB homologue being reported, which showed this enzyme was capable of repairing alkylation damage as would be expected (Simmons *et al*., 2012). Components of the BER pathway are encoded in trypanosomatid genomes, however homologues of XRCC1 and DNA ligase III have not yet been identified, suggesting short-patch BER may not occur or is significantly diverged from higher eukaryotes. Interestingly, *T. brucei* contain two distinct Pol β enzymes that are both active polymerases and dRP-lyases, and strikingly localise to the mitochondrion (Saxowsky *et al*., 2003). These two enzymes have been further characterised in *T. cruzi* and are suggested to have complementary roles in BER of kinetoplast (k)DNA (Lopes *et al*., 2008; Schamber-Reis *et al*., 2012). This example of otherwise nuclear eukaryotic proteins that are mitochondrial in trypanosomatids is frequently observed, and underlines the importance of maintaining kDNA in these protists. It is possible that Pol β may also be nuclear, as the *L. infantum* enzyme is nuclear (Taladriz *et al*., 2001) and shares a high degree of similarity with both the *T. brucei* and *T. cruzi* enzymes. It is notable that Pol β was the only X-family polymerase identified in the trypanosomatid genomes (El-Sayed *et al*., 2005). The majority of NER components are present in trypanosomatid genomes, however the mechanism of repair may differ slightly due to the duplication of some genes and the lack of others, exemplified by XP-A, of which no homologue was identified (El-Sayed *et al*., 2005). Also, given the previously mentioned absence of XRCC1 and DNA ligase III homologues, the ligation step during NER may have diverged significantly. It has been suggested that given the constitutive polycictronic transcription of protein-coding genes

in trypanosomes, that transcription-coupled NER may have an important role (Passos-Silva *et al*., 2010), although this is yet to be tested.

The MMR pathway is of particular relevance to *T. cruzi*, as it is thought to be an important source of antigenic diversity found within *T. cruzi* populations (reviewed in Machado *et al*., 2006). In *T. brucei* characterisation of MMR components thus far is consistent with this repair process operating (Bell *et al*., 2004). However, unlike *T. cruzi* it does not contribute to antigenic diversity, but does appear to have a regulatory role in HR, as following disruption of MMR components an increase in HR was observed, and this was regardless of whether DNA molecules were perfectly matched or contained mismatches (Bell and McCulloch, 2003). Double-strand break repair operates in *T. brucei*, however NHEJ appears to be absent. The trypanosomatid genomes lack key components such as DNA Ligase IV, XRCC4, and the majority of X-family polymerases (El-Sayed *et al*., 2005). Although Ku is present, DNA end-joining is not dependent upon this protein and is instead mediated by micro-homology (Burton *et al*., 2007). Glover *et al*., (2008) demonstrated that HR and micro-homology mediated end-joining repair the majority of chromosomal breaks in *T. brucei*, and observed no evidence for the presence of NHEJ. Ku has been implicated in telomere maintenance in *T. brucei* (Conway *et al*., 2002a). HR is particularly well characterised in *T. brucei* as these important pathogens use this DNA repair pathway in antigenic variation to escape their host's immune system. When in the mammalian host, the VSG protein coat of the trypanosome shields surface antigens from the host immune response. VSGs are expressed from a single allele, and HR is used to relocate VSG genes from sub-telomeric repositories to the promoter containing expression sites. By switching the VSG expressed the trypanosome can successfully escape the host's immune response (reviewed in Horn and McCulloch, 2010). A key eukaryotic HR component studied in *T. brucei* is the Rad51 recombinase; inactivation of this enzyme in cells impairs both HR and VSG switching (McCulloch and Barry, 1999). Additionally, a Rad51-independent HR pathway exists that is capable of recombining DNA with very short lengths of sequence homology, and this pathway also contributes to antigenic variation (McCulloch and Barry, 1999; Conway *et al*., 2002b). A key regulator of Rad51 is BRCA2 (breast cancer type 2 susceptibility protein); in *T. brucei* BRCA2 has undergone an expansion of BRC repeats, which are the motifs that facilitate interaction with Rad51, and these have been shown to be important for efficiency of HR and Rad51 localisation (Hartley and McCulloch, 2008). A recent report has suggested a dual role for BRCA2, in which this protein is required to maintain the VSG reporitories at sub-telomeric regions, and facilitate the re-localisation of Rad51 following DNA damage (Trenaman *et al*., 2012).

DNA damage tolerance remains largely uncharacterised in trypanosomatids. Central players in the regulation of TLS are present in trypanosomatid genomes, such as Rad6 and PCNA, in addition to a number of TLS polymerases. Genes encoding REV1, REV3 and 7 (Pol ζ), Pol η, and Pol κ are present, however, the Pol η paralogue Pol ι was not identified (El-Sayed *et al*., 2005). Stikingly, the *T. brucei* genome actually encodes 10 Pol κ's (El-Sayed *et al*., 2005). Only Pol η and κ have been initially characterised in *T. cruzi* (De Moura *et al*., 2009; Rajão *et al*., 2009). Pol η was a nuclear enzyme and although its ability to bypass UV photoproducts may be conserved, the authors suggest that in *T. cruzi* this enzyme is important for the cellular response to oxidative damage (De Moura *et al*., 2009). Similar conclusions were made for Pol κ also, however the Pol κ characterised (there are two in *T. cruzi*) was a mitochondrial enzyme (Rajão *et al*., 2009) and subsequent unpublished data from these authors suggest that the other Pol κ is nuclear (Passos-Silvia *et al*., 2010). Given the scarcity of active replication origins in these parasites (Tiengwe *et al*., 2012b), they may be more reliant on DNA repair and particularly damage tolerance mechanisms than other eukaryotes, and further characterisation of components of these pathways will no doubt prove insightful.

1.6. DNA Primases

A universal feature of DNA polymerases is their inability to initiate DNA synthesis *de novo*. Early studies by Arthur Kornberg and colleagues made observations to this effect, noting that although they had seen DNA polymerases extend already existing DNA chains, they had no evidence of these enzymes beginning the synthesis of a new DNA chain (Goulian and Kornberg, 1967; Goulian *et al*., 1968). They later implicated RNA polymerases in the initiation of DNA replication (Brutlag *et al*., 1971), which would synthesise short RNA chains to prime DNA synthesis (Wickner *et al*., 1972). In support of this, short RNA chains were found linked to nascent DNA in *E. coli* (Sugino *et al*., 1972). Subsequent work has showed that cellular DNA replication is absolutely dependant upon specialised DNA-dependant RNA polymerases that are distinct from classical RNA polymerases, and these enzymes were named primases (Scherzinger *et al*., 1977; Rowen and Kornberg, 1978).

1.6.1. Archaeal-eukaryotic primase (AEP) superfamily

DNA primases can be broadly divided into two superfamilies, those of the prokaryotic DnaG-like superfamily and those of the archaeo-eukaryotic primase (AEP)-like superfamily. Although functionally related, these two families have no evolutionary relationship (Avarind *et al*., 1998; Iyer *et al*., 2005) and are structurally distinct (Keck *et al*., 2000; Augustin *et al*., 2001). DnaG-like primases are required for the initiation of DNA

replication in bacteria and their phages and are functional as monomers, although often associating with the replicative DNA helicases. Whilst AEP-like primases are essential for the initiation of DNA replication in eukaryotes and archaea, and are usually heterodimeric enzymes composed of a small catalytic subunit and a large accessory subunit. In eukaryotes the heterodimeric primase exists in a complex with Pol α (Figure 1.13a) (reviewed in Frick and Richardson, 2001). Members of the AEP superfamily share homology with the small catalytic subunit (Prim1) of the eukaryotic heterodimeric primase, and robust sequence alignment methods have expanded this family to include homologous primases in viruses and bacteria (Aravind and Koonin, 2001; Weller and Doherty, 2001; Iyer *et al*., 2005).

1.6.2. Mechanism and structure of a DNA primase

The mechanism(s) of DNA primases remain poorly understood compared to other polymerases. This is largely due to the lack of high resolution stuctures, particularly in a ternary or quaternary complex with DNA and/or NTPs. The best-studied primases are those of the prokaryotic DnaG-like superfamily, with structures existing of these enzymes bound to single-stranded DNA (Corn *et al*., 2006) and NTPs (Rymer *et al*., 2012), in addition to apo structures (Keck *et al*., 2000; Kato *et al*., 2003). However, as initially suggested based on sequence analysis (Leipe *et al*., 1999), the topoisomerase/primase (TOPRIM) catalytic fold of DnaG-like primases is structurally distinct from that used by AEPs, and this was evident with the solving of the first structure of an AEP (Augustin *et al*., 1999; reviewed in Keck and Berger, 2000) (Figure 1.13b), therefore the mechanism of catalysis may differ between these two superfamilies.

Currently three structures exist of replicative AEPs: the initial apo (Augustin *et al*., 1999), bound to uridine 5'-triphosphate (UTP) (Ito *et al*., 2003), and as a heterodimeric complex (Lao-Sireix *et al*., 2005). However these are all archaeal enzymes, with no high resolution structure of a eukaryotic primase being solved to date. What is evident from these structures is that AEPs have a unique structure, and this common AEP catalytic fold has been observed in structures of divergent AEP homologues also (Lipps *et al*., 2004; Pitcher *et al*., 2007a). An extensive *in silico* analysis of the AEP superfamily by Iyer *et al*., (2005) has defined this shared catalytic core as consisting of two modules. The first is an amino-terminal $(\alpha\beta)_2$ unit that has no equivalent structure in the PDB database, and the second is a carboxyl-teriminal unit that, like the A- B- and Y-family DNA polymerases, is a highly derived RNA recognition motif (RRM). Three sequence motifs are highy conserved among the AEP catalytic fold: motif I is hhhDhD (where 'h' is a hydrophobic residue), motif II is sxH (where 's' is a small residue), and motif III is h-

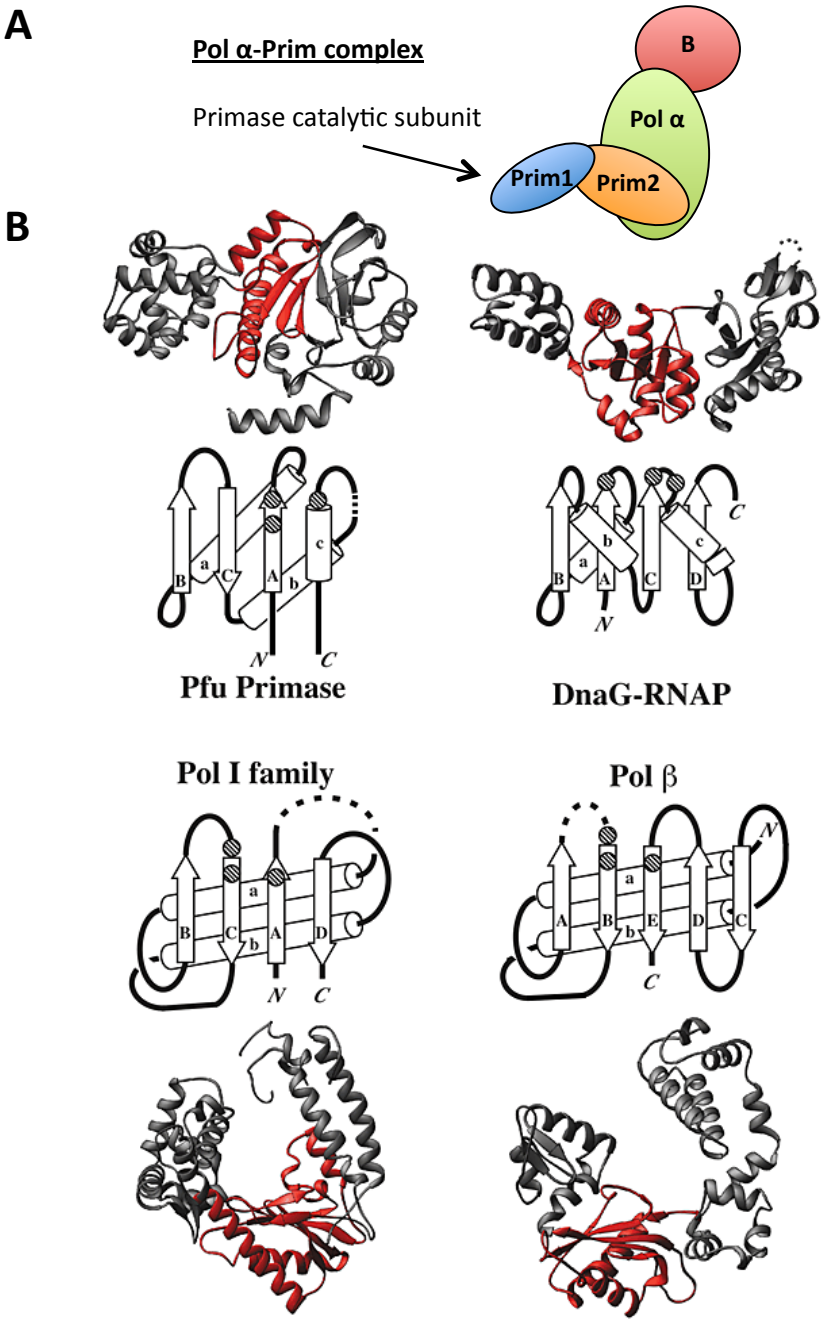


Figure 1.13. Structure of an archaeo-eukaryotic primase.
(**A**) Composition and subunit contacts of the eukaryotic Pol α-Prim complex. The eukaryotic DNA primase is a heterodimer composed of a catalytic (Prim1/PriS) and accessory (Prim2/PriL) subunit that exists associated with the Pol α heterodimer composed of a catalytic (Pol α) and accessory (B) subunit. Prim1/PriS is the reference member of the AEP superfamily. (**B**) Comparison of the polymerase domain structure of *P. furiosus* primase (PriS; AEP) (Augustin *et al*., 2001), *E. coli* DnaG primase RNA polymerase domain (RNAP) (Keck *et al*., 2000; Podobnik *et al*., 2000), and canonical DNA polymerases *E. coli* Pol I (Ollis *et* al., 1985) and rat Pol β (Pelletier *et al*., 1994; Davies *et al*., 1994). Active site regions are highlighted in red and drawn as topology diagrams, with the conserved catalytic triads indicated as hatched circles. The structures of the AEP and DnaG primase are clearly distinct, and neither contain the characteristic fingers and thumb sub-domains of canonical DNA polymerases. Despite the difference in structure it is notable that the catalytic triad of the *Pfu* primase and the canonical DNA primases are arranged in a similar manner. Figure taken from Keck and Berger, 2001.

(where '-' is an acidic residue). Various structural and site-directed mutagenesis studies of AEPs have shown that these three motifs are essential for catalysis, with residues in motif I and III being required for binding of a divalent metal ion and motif II required for nucleotide binding (Copeland and Tan, 1995; Augustin *et al*., 2001; Ito *et al*., 2003; Lao-Sirieix and Bell, 2004). These catalytic motifs reside between the two modules of the AEP catalytic fold. Despite being a unique structure, the catalytic sites of AEPs and some DNA polymerases, namely the X-family Pol β, are super-imposable (Augustin *et al*., 2000), which is consistent with the homology between AEPs and X-family polymerases first discovered by Kirk and Kuchta (1999). This architectural similarity, in addition to the requirement for divalent metal ions for catalysis, suggests that AEPs use the two-metal ion dependant mechanism for elongation, like DNA polymerases (Figure 1.4) (Steitz *et al*., 1994). This is supported by a recent structure of a divergent bacterial AEP, the NHEJ polymerase PolDom, in which this AEP was crystallised bound to a DNA end, UTP, and two divalent metal ions, notably without a primer strand provding the 3' hydroxyl group. The binding of both metal ions rendered the catalytic centre of this enzyme ready for phosphoryl transfer (Brissett *et al*., 2011). Interestingly, a two-metal ion dependant mechanism of catalysis has also been suggested for the structurally distinct bacterial DnaG primase (Rymer *et al*., 2012). However, for a complete understanding of AEPs, in particular eukaryotic AEPs and the mechanism of initiation, further high resolution structures are required.

There is a minimum of five discrete steps proposed for oligoribonucleotide synthesis by a DNA primase (reviewed in Frick and Richardson, 2001; Arezi and Kuchta, 2000). First, the primase binds to a single-stranded DNA template, then, as primer synthesis does not occur randomly, the enzyme slides along the DNA until the initiation site is located. The eukaryotic primases require a minimal initiation sequence compared to their archaeal and bacterial counterparts, requiring a templated pyrimidine to code the 5' terminal nucleotide of the primer (Davey *et al*., 1990; Grosse and Krauss, 1985; Tseng and Ahlem, 1984). Once at the initiation sequence the primase sequentially binds two NTPs, with one proposed to bind the so-called initiation site and a second bind the active site (Frick *et al*., 1999). The initiation site binds the second NTP providing the 3' hydroxyl to attack the α-phosphate of the first NTP that binds to the active site. Thus the first NTP bound becomes the second NTP of the primer (Sheaff and Kuchta, 1993). With a quaternary structure formed of enzyme-DNA-NTP-NTP, the primase catalyses dinucleotide formation. This occurs in the unorthodox direction of 3'-5', and results in the release of an inorganic pyrophosphate (Sheaff and Kuchta, 1993). The primer is then

extended at the 3' end with NTPs until hand-off to a DNA polymerase (e.g. Pol α), which utilises the 3' hydroxyl group and subsequently extends with dNTPs.

1.6.3. Eukaryotic AEPs – priming DNA replication

The eukaryotic replicative DNA primase (Prim1) is conserved from the earliest diverging eukaryotes to man, and in all organisms examined thus far, has existed as a four-subunit complex with Pol α, and every subunit of this complex is essential for cell viability (Lucchini *et al*., 1987; Foiani *et al*., 1989; Sugino, 1995; Muzi-Falconi *et al*., 2003). This underlines the critical role of the Pol α-Prim complex in DNA replication, for the initiation of replication at origins and each Okazaki fragment (see section 1.2.2). Of the four-subunit complex, the catalytic DNA primase Prim1 subunit is the smallest (50 kDa in humans), and is capable of primer synthesis on its own but is very unstable, requiring its accessory Prim2 subunit (p58) (Schneider *et al*., 1998). In addition to being required for efficient primer synthesis, the Prim2 subunit is also required for functional interaction with the catalytic Pol α subunit (Longhese *et al*., 1993). A recent study demonstrated an important role for Prim2 in loading the Pol α-Prim complex on RPA-coated single stranded DNA, through specific interactions between a highly conserved carboxyl-terminal domain of Prim2 with RPA (Vaithiyalingam *et al*., 2010). The Prim2 subunit also contains a nuclear localisation signal that is required for the correct localisation of the catalytic subunit, through a so-called "piggy-back" mechanism (Mizuno *et al*., 1996). As would be expected for an essential enzyme, Pol α-Prim is under strict regulation (reviewed in Muzi-Falconi *et al*., 2003). The mammalian enzyme is phosphorylated in a cell cycle dependant manner (Nasheuer *et al*., 1991), and this has been shown to effect catalytic activity, with maximum activity observed when the B-subunit of Pol α is phosphorylated but the catalytic Pol α subunit is not (Schub *et al*., 2001).

In addition to its central role in DNA replication, Pol α-Prim is also inherently important for the cellular response to DNA damage, as continued primer synthesis at stalled replication forks is critical for activation of the intra-S-checkpoint (Micheal *et al*., 2000; MacDougall *et al*., 2007; Van *et al*., 2010). In addition it has been shown to be both necessary and sufficient for ubiquitinated-PCNA mediated DNA damage tolerance (Chang *et al*., 2006; reviewed in Chang and Kimprich 2009). Work in yeast using primase mutants suggested that Prim1 was involved in coupling DNA replication to the DNA damage response, through a pathway dependant upon the ATM and Chk2 kinases (Marini *et al*., 1997). The authors speculated that regulation of Pol α-Prim is required to prevent re-priming after replication-fork blocking lesions, thus slowing replication and providing time for repair (reviewed in Foiani *et al*., 1997). The chromatin remodeler

BAF180 was recently implicating in facilitating this re-priming in yeast and mammalian cells (Niimi *et al*., 2012). Despite numerous studies regarding re-priming in eukaryotic cells (discussed previously in section 1.4.5), no mechanistic demonstration has been reported to date. However, a mechanism for re-priming has been delineated using bacterial proteins. Heller and Marians (2006) demonstrated that when the 3' hydroxyl group of the nascent leading strand is unavailable, a single replicative helicase (DnaB) complex is sufficient to co-ordinate DnaG-dependant priming downstream on both the leading and lagging strand. Subsequent work has shown this discontinous leading strand synthesis occurs on a plasmid containing a leading strand CPD *in vitro*, and that this process was not dependant upon a replication restart machinery but the DnaG primase, demonstrating it is an inherent feature of the bacterial replisome (Yeeles and Marians, 2011).

In addition to the essential replicative DNA primase of the Pol α-Prim complex, two recent studies have identified novel AEPs present in trypanosomatid genomes, with the *T. brucei* homologues of these AEPs being initially characterised (Hines and Ray, 2010 and 2011). These two enzymes were confusingly named Pri1 and Pri2 and both localised to the mitochondrion, and so will be referred in this thesis kinetoplast (k)Pri1 and kPri2. Both were demonstrated to be active DNA primases *in vitro*, capable of *de novo* DNA-dependant RNA synthesis to produce RNA primers that could be subsequently extended by a DNA polymerase. kPri1 and kPri2 were both reported to be essential for cell growth, suggestive of a role in kDNA replication. The mitochondrial genome of *T. brucei* consists of a few thousand so-called minicircles (~1 kb) and a few dozen maxicircles (~23 kb), catenated into a chainmail like structure that *in vivo* forms a kDNA disk (reviewed in Liu *et al*., 2005). kPri1 was suggested to be primarily responsible for initiating DNA synthesis of the maxicircles, whilst kPri2 was suggested to be responsible for intiating minicircle DNA replication (Hines and Ray, 2010 and 2012).

1.6.4. Archaeal AEPs – versatile primase-polymerases?

The archaeal AEPs consisting of a small catalytic (PriS) and a large accessory (PriL) subunit, provide the only high resolution structures of replicative AEPs as discussed previously, and possess some novel activities (Figure 1.14). Initial studies of the replicative AEP from a *Pyrococcus* species demonstrated this enzyme was capable of initiating DNA synthesis with dNTPs and synthesising fragments up to 6 kilo bases in length, although it was incapable of generating RNA primers (Bocquier *et al*., 2001). However, a subsequent study showed that when this enzyme functioned as a heterodimer rather than a monomer, the length of DNA chains synthesised were greatly

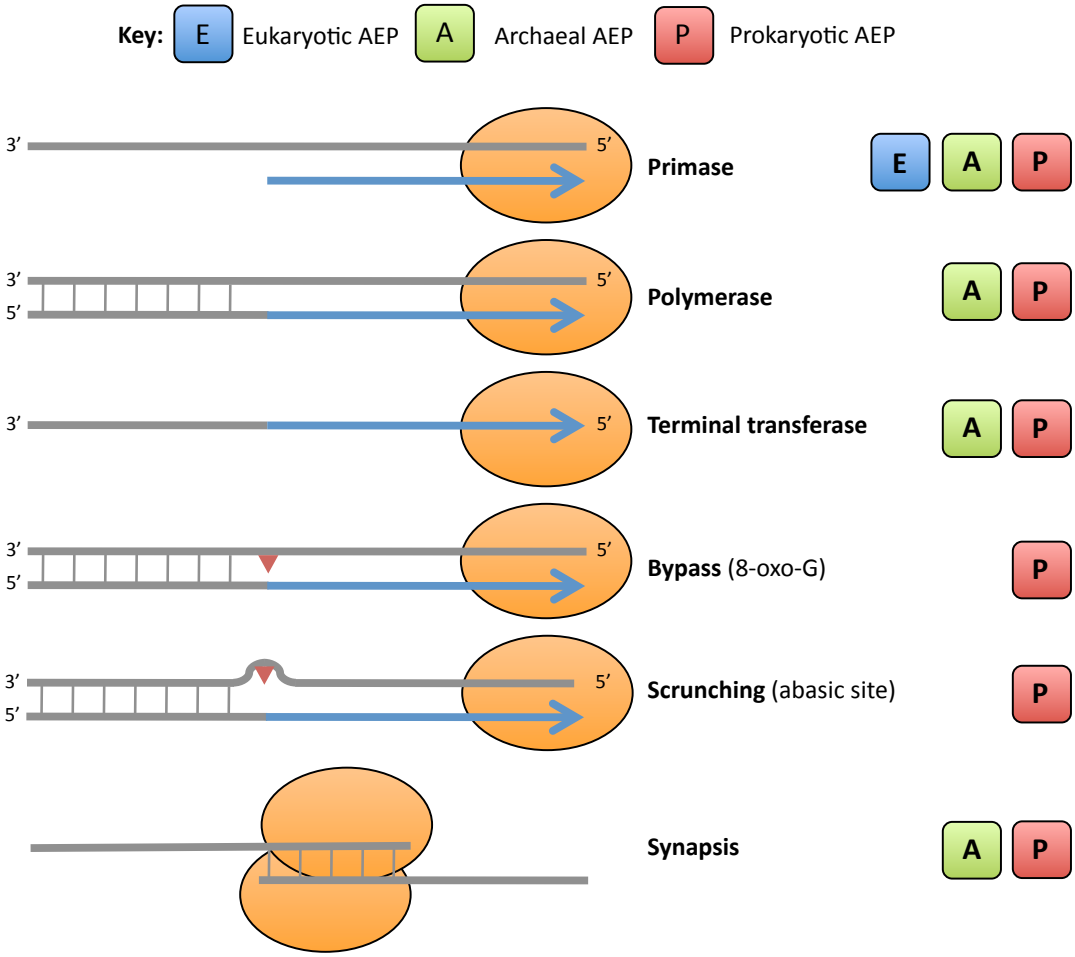


Figure 1.14. The promiscuous activities of archaeal-eukaryotic primases.
The archaeo-eukaryotic primase (AEP) superfamily contains a number of versatile nucleotidyl transferases. The extent of reported activities are depicted above; AEP enzyme (orange circle), nascent DNA/RNA (blue arrow), minor helix-distorting DNA lesions (red triangle). Coloured boxes indicate whether the activity was observed from a eukaryotic (blue), archaeal (green), or prokaryotic (red) AEP enzyme. Prim1, the catalytic primase component of the Pol α-Prim complex, is capable of *de novo* DNA-dependant RNA polymerase activity i.e. it can prime using nucleoside triphosphates (NTPs) to generate an RNA primer (reviewed in Frick and Richardson, 2001). The archaeal Prim1 homologue PriS, which exists as a heterodimer with it's large accessory subunit PriL, can catalyse *de novo* DNA-dependant RNA/DNA polymerase activity, priming with either NTPs or deoxy (d)NTPs to make RNA or DNA primers respectively. PriSL is also a competent DNA-dependant DNA/RNA polymerase and has been shown to possess template-independent terminal transferase activity (reviewed in Lao-Sirieix *et al.*, 2005a). The AEP domain of Ligase D (PolDom) from bacteria has been reported to catalyse all of the above activities with the exception of making *de novo* DNA primers. Also PolDom is capable of incorporating and extending opposite 8-oxo-guanine lesions and bypassing abasic sites by dissociating and re-aligning (scrunching) the DNA template (reviewed in Pitcher *et al.*, 2007b). PolDom has been demonstrated to facilitate DNA end-synapsis also, using microhomology, (reviewed in Brissett and Doherty, 2009), which is required during DNA break repair, and this was recently suggested for the archaeal PriSL also (Hu *et al.*, 2011).

reduced, and the enzyme was able to generate RNA primers (Liu *et al*., 2001). No such versatility of the eukaryotic primase has been reported to date. This activity was not restricted to the *Pyrococcus* enzyme as a *Sulfolobus* enzyme was also demonstrated to initiate DNA synthesis with both NTPs and dNTPs, and synthesise chains of up to 1 and 7 kb respectively (Lao-Sirieix and Bell, 2004), suggesting this may be a conserved feature of archaeal AEPs. A question remains as to whether there is an *in vivo* function to initiating strand synthesis with dNTPs, and current data suggests against it: archaeal Okazaki fragments are known to contain a 5' RNA chain (Matsunaga *et al*., 2003; Beattie and Bell, 2012) and the archeal enzymes have a much greater affinity for NTPs (Lao-Sirieix and Bell, 2004). Another striking activity of these enzymes is there ability to catalyse terminal transferase (Lao-Sirieix and Bell, 2004; De Falco *et al*., 2004), providing yet another link between the family-X polymerases and AEPs. Additionally, a recent study suggested through *in vitro* experiments that the archaeal AEP can bridge non-complementary DNA ends and facilitate repair of double-strand breaks (Hu *et al*., 2011), however no *in vivo* work has verified this model.

A novel AEP-like enzyme has been characterised in another *Sulfolobus* species encoded on the pRN1 plasmid. This is a multidomain protein with ATPase, primase, and polymerase activities (Lipps *et al*., 2003). The AEP-like domain of this enzyme, referred to as pRN1-PrimPol, is another example of the versatility of an AEP. Although only accepting dNTPs for elongation, this enzyme could initate DNA synthesis using a single NTP and extend with dNTPs, and further extend the subsequent RNA-DNA primer over several kilo bases (Beck and Lipps, 2007). The crystal structure of this enzyme confirmed its polymerase domain was a member of the AEP superfamily, and it shared significant structural similarity to other archaeal AEPs (Lipps *et al*., 2004)

1.6.5. Bacterial AEPs – NHEJ polymerases

One of the first hints of an alternate function for an AEP was the identification of AEP homologues in bacteria, as bacteria already contain a dedicated DnaG replicative primase. Notably, the AEP homologues were often part of a multidomain protein called Ligase D (LigD), consisting of putative DNA ligase and nuclease domains, encoded by a gene that was co-operonic with homologues of the eukaryotic Ku DNA repair protein (Koonin *et al*., 2000; Aravind and Koonin, 2001; Doherty *et al*., 2001; Weller and Doherty, 2001). Ku is an essential component of the NHEJ double-strand break repair pathway in eukaryotes, required for recognition and synapsis of broken DNA ends and initiating the recruitment of repair enzymes (reviewed Brissett and Doherty, 2009). Given that co-operonic genes in bacteria can often function in a common pathway (Dandekar *et al*.,

1998), this inferred the existence of NHEJ in prokaryotes (Koonin *et al*., 2000; Aravind and Koonin, 2001; Doherty *et al*., 2001; Weller and Doherty, 2001). Shortly after, a physical and functional interaction between Ku and LigD in the repair of double-strand breaks was demonstrated (Weller *et al*., 2002), and further, it was demonstrated that just these two proteins comprise a minimal two-component NHEJ repair machine in bacteria (Della *et al*., 2004). This is largely due to the versatility of LigD, which provides the entire DNA end processing activities (polymerase, nuclease, and ligase) required to repair a break (Della *et al*., 2004). The AEP polymerase domain (PolDom) of LigD was demonstrated to be a particularly versatile polymerase, with this single domain being capable of catalysing a variety of nucleotidyl transferase activites (Figure 1.14). PolDom could perform template-dependant DNA/RNA polymerase and gap-filling activity, template-dependant RNA-priming activity, and template-independent terminal transferase activity on both single-stranded and blunt ended double-stranded DNA (Della *et al*., 2004; Pitcher *et al*., 2005, 2007a). Additionally, PolDom could perform error-free gap-filling opposite 8-oxo-G containing templates (Pitcher *et al*., 2007a). PolDom was also particularly specialised at recognising DNA breaks, with a major determinant being the 5' phosphate moiety (Pitcher *et al*., 2007a), and has even been shown to mediate DNA end-synapsis using microhomology (Brissett *et al*., 2007). In summary, PolDom possesses all the DNA/RNA synthetic processes that could be required at a DNA break, which in eukaroytes is divided between three family-X DNA polymerases. A notable feature of PolDom was its marked preference of incorporating NTPs instead of dNTPs, which *in vivo* is suspected to be required when cellular pools of dNTPs are low, as in the case of stationary phase bacteria (Della *et al*., 2004; Pitcher *et al*., 2007a). Additionally, the bacterial LigD is co-opted by mycobacteriphage encoded Ku to facilitate the circularisation of their genome (Pitcher *et al*., 2006). Various structures exist of PolDom (Zhu *et al*., 2006; Pitcher *et al*., 2007a) confirming it is an AEP polymerase and is therefore the only AEP with a proven cellular role outside of DNA replication.

1.7. PrimPol, a novel eukaryotic primase-polymerase

The Doherty laboratory has a long-standing interest in AEP-like enzymes, demonstrated by the extensive characterisation of the bacterial NHEJ primase-polymerase PolDom, as detailed in the previous section (reviewed in Pitcher *et al*., 2007b; Brissett and Doherty, 2009). The identification of a variety of novel AEPs in bacteria prompted the question of whether additional uncharacterised AEPs existed in eukaryotes. Iterative searching of the available eukaryotic genomes for genes with sequence homology to AEPs identified the human gene *CCDC111* (alternate name FLJ33167), which is located on chromosome 4 (4q35.1) and encodes coiled-coil domain containing protein (CCDC)111. This gene was

also identified by Iyer and colleagues to be a member of the AEP superfamily based on *in silico* analysis (referring to the protein an EukPrim2 for eukaryotic primase 2), specifically a member of the nucleo-cytoplasmic large DNA virus (NCLDV)-herpesvirus primase clade (Iyer *et al*., 2005), but to date has remained uncharacterised. Preliminary characterisation of CCDC111 in the Doherty lab and collaborators Luis Blanco's (CBMSO, Madrid) demonstrated that the recombinant human protein was a versatile nucleotidyl transferase *in vitro*, capable of template-dependant primase and polymerase activities with both NTPs and dNTPs, resembling the catalytic activities of the versatile bacterial and archaeal AEP enzymes. Thus, in keeping with primase nomeculature, CCDC111 was renamed PrimPol (primase-polymerase) to reflect its intrinsic enzymatic activities.

As reported by Iyer and colleagues (2005), iterative BLAST searching with human PrimPol and putative homologues indicates PrimPol is present in a broad range of unicellular and multicellular eukaryotes including animals, plants, and protists, but is not present in prokaryotic and archaeal genomes. PrimPol is not conserved throughout all eukaryotes however, notably absent from *Drosophila, Caenorhabditis elegans* and all but one fungus, the parasitic *Batrachochytrium dendrobatidis*. Iyer and colleagues suggest that the unusual distribution of PrimPol homologues suggests that the gene for this protein was acquired early in eukaryotic evolution and then lost independently on multiple occasions in fungi and animals (Iyer *et al*., 2005). Alignment of putative PrimPol homologues reveals several conserved regions, principally divided into two domains; the first is a catalytic AEP domain towards the amino-terminus thus identifying PrimPol as a member of the AEP superfamily (Figure 1.15 and 1.16). The PrimPol AEP domain contains the three catalytic motifs conserved in all AEP-like enzymes (Figure 1.15) (see section 1.6.2) (Iyer *et al*., 2005). Motif I in PrimPol homologues has the consensus LYFDLE with invariant DxE residues, which is unusual among members of the AEP superfamily as motif I usually contains invariant DxD with the exception of the pRN1-PrimPol (Lipps *et al*., 2003). Motif II in PrimPol homologues is an invariant SxH, and motif III is an invariant xD. Based on previous studies of AEPs (Copeland and Tan, 1995; Augustin *et al*., 2001; Ito *et al*., 2003; Lao-Sirieix and Bell, 2004) residues in motif I and III would be predicted to be required for binding of a divalent metal ions and motif II required for nucleotide binding. The second conserved region among PrimPol homologues is a carboxyl-terminal invariant CHC_2 motif with homology to the human herpesvirus (HHV) UL52 primase (Figures 1.15 and 1.16), and is predicted to be a highly derived zinc-finger/ribbon (Iyer *et al*., 2005). Zinc binding domains are a common feature among primases and often play a number of critical roles in primer synthesis (reviewed in

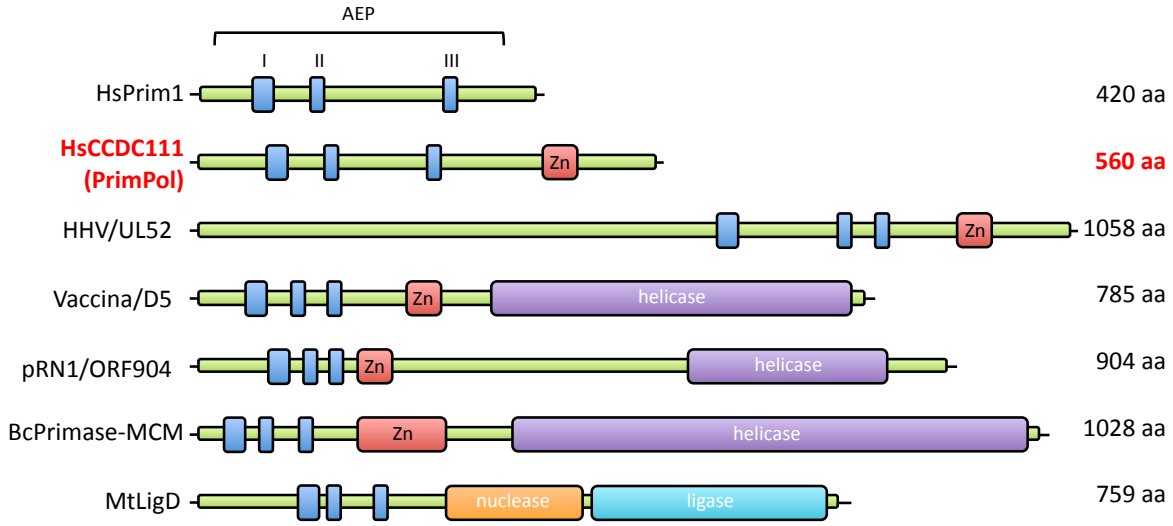


Figure 1.15. Domain organisation of various archaeo-eukaryotic primases.
Domain organisation of various AEPs are depicted: motifs I, II, and III, of the catalytic AEP domain (blue boxes), zinc finger (Zn) motifs (red boxes), and additional domains often associated with AEP enzymes such as helicase (purple boxes), nuclease (orange box) and ligase (cyan box) domains are shown. The name of each AEP enzyme and length in amino acids is also indicated. Prim1 is the reference member of the AEP superfamily and together with the novel AEP coiled-coil domain containing protein (CCDC)111, are the only AEPs identified to date in the human genome. The human herpesviral (HHV) UL52 primase and vaccina D5 primase are viral AEPs, whilst ORF904 is encoded on the pRN1 plasmid of the archaeon *Sulfolobus*. The prokaryote *Bacillus cereus* genome encodes an AEP-helicase fusion and a number of bacteria, including *Mycobacterium tuberculosis*, encode the NHEJ primase Ligase D (LigD). Domain organisations were deduced from Iyer *et al*., 2005.

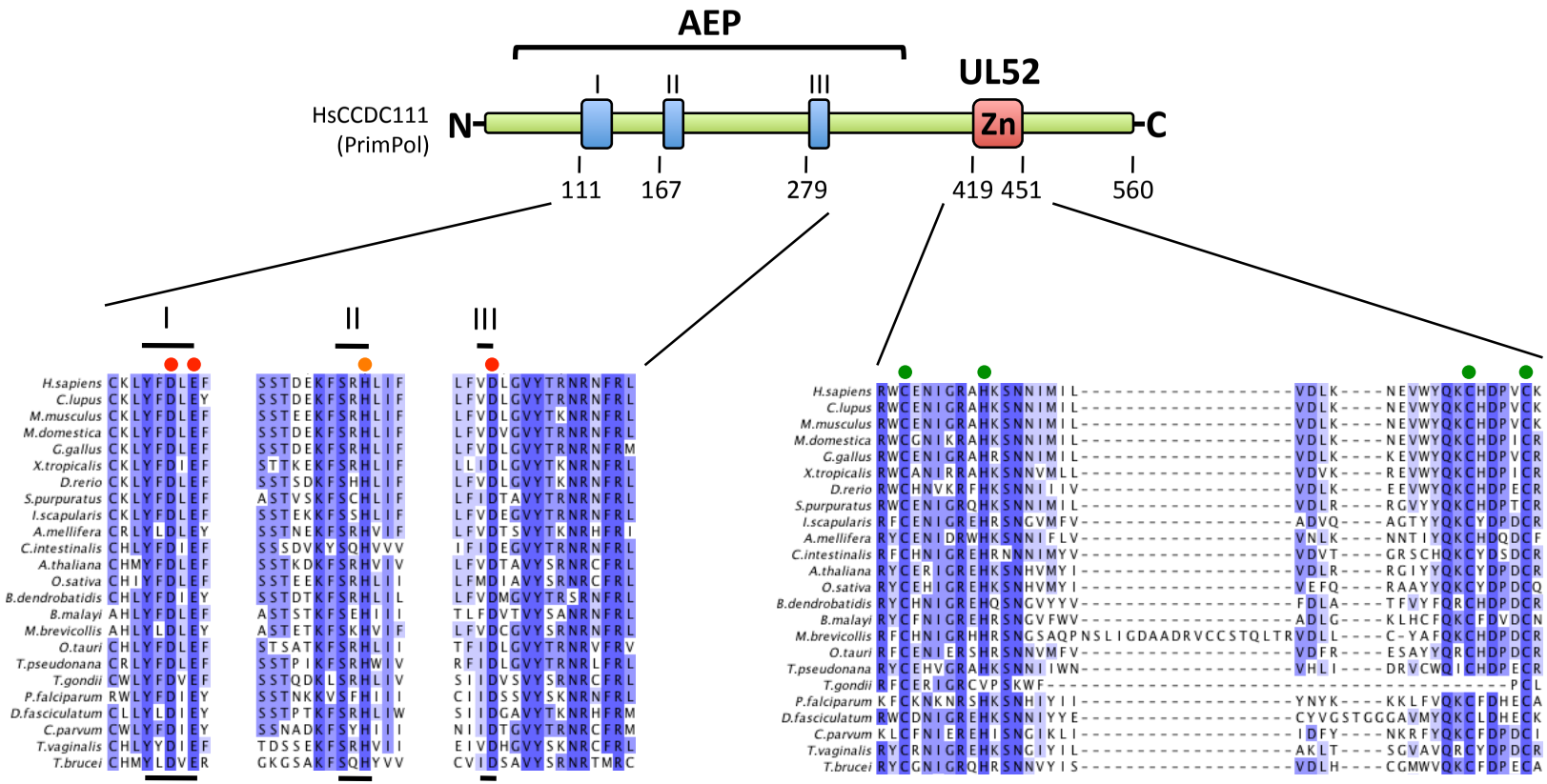


Figure 1.16. Conserved domain of the PrimPol (CCDC111) famliy.
PrimPol contains two identifiable domains. Towards the amino-terminus there is an archaeo-eukaryotic primase (AEP) domain containing the three motifs predicted to be essential for catalysis (blue boxes). Residues in motif I and III are predicted to be the catalytic carboxylates required for binding of divalent metal ions (red circles), whilst residues in motif II are required for nucleoside triphosphate (NTP) binding (orange circle). Towards the carboxyl-terminus is a CHC_2 zinc finger motif (Zn) (green circles) with homology to the herpesviral UL52 primase. Numbers on the domain schematic refer to the amino acid number for the beginning of these motifs in human PrimPol. The multiple sequence alignment was generated using PrimPol homologues from a broad range of multicellular and unicellular eukaryotes, which were identified by iterative PSI-BLAST searches. Alignments were generated using ClustalW2 (Larkin *et al.*, 2007), TCoffee (Notredame *et al.*, 2000), MUSCLE (Edgar, 2004) and edited in JALview (Waterhouse *et al*, 2009). Blue shading indicates ≥40 % sequence identity. See appendix for full sequence alignment and accession numbers of PrimPol homologues used.

Kutcha and Stengel, 2010). The zinc finger of the HHV UL52 primase has been implicated in DNA binding, primase activity, and for correct functioning of the heterotrimeric HHV helicase-primase complex (Biswas and Weller, 1999; Chen *et al*., 2001).

The overall aim of this thesis was to determine the cellular role of PrimPol. Two approaches were used, the first was to characterise a PrimPol homologue in a relatively 'simple' eukaryotic model organism using both cellular and biochemical methods (Chapters 3 and 4). The second approach was to characterise PrimPol in cultured human cells, to first determine its localisation within the cell (Chapter 5) and then build on this information, and other research from the Doherty laboratory, to uncover the specific function(s) of this novel polymerase in human cells (Chapter 6). Attempts were also made to identify interacting partners of human PrimPol, with the aim of shedding light on the cellular pathways it may operate in (Chapter 7).

Chapter 2

Materials and Methods

All chemicals were purchased from Sigma Aldrich or Fisher Scientific unless stated.

2.1. Preparation of plasmid DNA

2.1.1. Preparation of competent *E. coli* DH5α

The *E. coli* strain DH5α was used to prepare DNA plasmids. Three millilitres lysogeny broth (also called Luria-Bertani, LB) medium (1 % (w/v) tryptone, 0.5 % (w/v) yeast extract, 1 % (w/v) NaCl, pH 7) was inoculated with a single colony of DH5α and incubated overnight with shaking at 37 ºC. The saturated culture was then diluted in 250 ml super optimal broth (SOB) medium (2 % (w/v) tryptone, 0.5 % (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgSO_4$, 10 mM $MgCl_2$) and incubated at 18 ºC in a baffled flask with shaking until the optical density measured at 600 nm (OD_{600}) was ~0.4. Following 10 minutes cooling on ice, cells were collected by centrifugation (4000 rpm, 10 minutes, 4 ºC), resuspended in 80 ml ice-cold transformation buffer (100 mM PIPES pH 6.7, 15 mM $CaCl_2$, 250 mM KCl, 55 mM $MnCl_2$) and incubated on ice for a further 10 minutes. Cells were again collected by centrifugation and gently resuspended in 20 ml ice-cold transformation buffer. Dimethyl sulfoxide (DMSO) was added to a final concentration of 7 % and following 10 minutes on ice, cells were divided into 50 – 200 µl aliquots and frozen in liquid nitrogen to be stored at -80 ºC.

2.1.2. Transformation of competent DH5α

To 50 µl chemically competent DH5α (section 2.1.1), which had been thawed on ice, typically 1 µl plasmid DNA from a MiniPrep (~100 ng) (section 2.1.3) or 3-5 µl ligation reaction (section 2.1.4) was added. The cell and DNA mix was incubated on ice for 10 minutes and then heat shocked at 42 ºC for 1 minute before being incubated on ice for a further 15 minutes. To the mixture, 1 ml LB medium was added before incubation at 37 ºC with shaking for ~40 minutes. One hundred microliters was plated on LB agar plates (LB medium solidified with 1.5 % (w/v) agar) containing the appropriate antibiotics and incubated overnight at 37 ºC to allow colonies to grow. Ampicillin was used at 100 µg/ml and kanamycin at 30 µg/ml final concentrations.

2.1.3. Plasmid DNA amplification and purification

Plasmid DNA was transformed into DH5α (section 2.1.2) and a single colony selected to inoculate 3 ml of LB medium containing the appropriate antibiotic (100 µg/ml ampicillin or 30 µg/ml kanamycin). Following overnight incubation at 37 ºC with shaking, the DNA was prepared using the QIAprep Spin MiniPrep Kit (Qiagen) according to the manufacturer's instructions and stored at -20 ºC. Plasmid DNA was typically eluted in 50

µl providing a final concentration of 100 ng/µl. If a larger quantity of DNA was required, the 3 ml starter culture was incubated for ~8 hours before using 0.5-1 ml to inoculate 50-100 ml LB medium with the appropriate antibiotic. Following an overnight incubation at 37 ºC with shaking, the DNA was prepared using the QIAfilter Plasmid Midi Kit (Qiagen) according to the manufacturer's instructions. Plasmid DNA was typically eluted in 50 µl providing a final concentration of ~2-3 µg/µl. The yield was determined using a NanoDrop spectrophotometer (Thermo Scientific) and the DNA stored at -20ºC.

2.1.4. Agarose gel electrophoresis of DNA

DNA was typically resolved on a 1 % (w/v) agarose TAE (0.4 M Tris-Acetate pH 8, 1 mM EDTA) gel containing ~0.3 µg/ml ethidium bromide and electrophoresed at 100 V for ~20 minutes. Samples were loaded in DNA loading buffer (2.5 % (w/v) Ficoll 400, 11 mM EDTA, 3.3 mM Tris pH 8, 0.017 % (w/v) SDS, 0.015% (w/v) bromophenol blue), which was supplied in a 6x stock (New England Biolabs), and resolved alongside a 1 kb DNA ladder (New England Biolabs). DNA was visualised using a UV illuminator (Syngene InGenius Gel Documentation System) and images analysed using GeneSnap (Syngene).

2.2. Molecular cloning

2.2.1. Polymerase chain reaction

Primers for the polymerase chain reaction (PCR) were designed typically to have a Tm of ~55-65 ºC, with the T_m difference being limited to 2-3 ºC between primers, using the modified Breslauer's method (Breslauer *et al*, 1986) with the Finnzymes Tm calculator (Thermo Scientific). The PCR was performed using Phusion high-fidelity DNA polymerase (Thermo Scientific) according to the manufacturer's instructions. Reactions of 50 µl were assembled in 0.2 ml tubes containing typically ~10 ng template plasmid or ~200 ng genomic DNA, 0.5 µM forward and reverse primers, and 1 U Phusion. In a Techne TC3000G thermo-cycler the reaction was initially denatured for 3 minutes at 98 ºC before 25 cycles of 98 ºC for 15 seconds, annealing temperature for 15 seconds, and elongation at 72 ºC, and a final elongation of 72 ºC for 2 minutes. The annealing temperature was typically 3 ºC above the Tm's of the primers, and the elongation time was calculated according to the desired amplicon's size, with the polymerase's synthesis rate of 1 kb every 15-30 seconds. Optimisation of the PCR reaction often entailed trying a range of annealing temperatures and/or inclusion of DMSO or extra $MgCl_2$ into the reaction. To determine whether the PCR was successful, 5 µl of the product was resolved by agarose gel electrophoresis (section 2.1.4). The remaining PCR product was purified using the QIAquick PCR purification kit (Qiagen) according to the manufacturer's protocol.

2.2.2. Mutagenesis and inverse deletion PCR

Primers were designed for site-directed mutagenesis using either the QuickChange protocol (from Stratagene), which uses a completely complementary primer pair containing the mutation, or a modified version of the QuickChange protocol from Zheng *et al*. (2004), in which partially overlapping primers are used. The latter protocol minimises primer dimerisation to ensure primer-template annealing is more favourable than primer-primer annealing during the PCR. Phusion high fidelity DNA polymerase (Thermo Scientific) was used for the reactions, with the contents of the reaction and the program being the same as previously described (section 2.2.2), except the final elongation at 72 ºC was 10 minutes. Optimisation involved titration of template (0.1-100 ng) and primers (0.2-2 µM) as well as inclusion of DMSO and $MgCl_2$. If the primer Tm's were above 72 ºC then a two-step PCR was performed, in which the annealing step was omitted. As a control, a mutagenesis PCR reaction was setup omitting Phusion DNA polymerase. The successful PCR product (ideally visible when resolved by agarose gel electrophoresis), along with the 'no phusion' control, were subjected to the QIAquick PCR purification kit (Qiagen) and digested with an excess of 10-20 U DpnI (New England Biolabs) in a 30 µl reaction for ~3 hours at 37 ºC. One microliter of the reaction products were transformed into DH5α and around a third of the transformation plated and incubated overnight at 37 ºC. Colonies were screened by digest (section 2.2.3) if a restriction site was also inserted, or otherwise checked by sequencing (section 2.2.5).

Inverse PCR requires a primer pair in a "back-to-back" orientation, being the opposite of a usual PCR. By having a phosphate present on the 5' end of these primers, allowing ligation of these ends, this strategy can be used to remove sequences from a construct by having the primers flanking the area to be deleted (Ochman *et al*., 1988; Hemsley *et al*., 1999). Following the PCR reaction, the product was treated with DpnI and then an overnight ligation was set up at 12 ºC using 100 ng PCR product and 1 µl (400 U) T4 DNA Ligase (with a no DNA ligase control). The products were transformed into DH5α (section 2.1.2) and the subsequent colonies screened by restriction digest (2.2.3).

2.2.3. Restriction digest

When using restriction endonucleases in generating a construct, typically 20 µl of MiniPrep plasmid DNA (~2 µg) or 20 µl of purified PCR product were digested in a 40 µl reaction with an excess of ~20 U of the appropriate restriction enzyme(s) and reaction buffer (New England Biolabs). Following incubation at 37 ºC for between 1 and 16 hours the digested plasmid or PCR product were purified using either the QIAquick PCR purification kit or first resolved by agarose gel electrophoresis (section 2.1.4) and purified

from an excised gel slice using the Geneclean II kit (Qbiogene Inc.), both according to manufacturer's instructions.

Restriction digestion was also used to screen clones when generating a construct. The sequence of the theoretical construct was analysed using the NEBcutter web tool (Vincze *et al*, 2003) and restriction enzyme(s) chosen to identify the correctly made construct. Digests were performed typically with 3 µl of miniprep plasmid DNA in a final volume of 10 µl with ~10 U restriction endonuclease. Half of the reaction was resolved by agarose gel electrophoresis (section 2.1.4) and compared to the digested parental plasmid DNA and the theoretical digest from NEBcutter, to determine if the clone was correct.

2.2.4. Ligation

T4 DNA Ligase (New England Biolabs) was used to ligate restriction digested plasmid DNA and PCR products when generating a construct. The concentration of the digested DNA was determined using a NanoDrop spectrophotometer (Thermo Scientific) and the ligation reaction was set up with an insert (I) to vector (V) ratio of 6:1 using the equation $I_{(ng)} = 6x\ [I_{(bp)} / V_{(bp)}] \times V_{(ng)}$. Two hundred units (0.5 µl) of T4 DNA Ligase was used, in a final volume of 20 µl with diluted T4 DNA Ligase buffer (New England Biolabs). The ligation was incubated at room temperature for 1 hour before 3-5 µl was transformed into DH5α (section 2.1.2), the remainder of the ligation was incubated at 12-16 ºC overnight and transformed if required, otherwise stored at -20 ºC.

2.2.5. Sequencing

Sequencing of plasmid DNA or PCR products was performed by GATC biotech using universal primers or gene specific primers (Tables 2.4, 2.6, 2.9, 2.11). The sequencing chromatogram was read using Chromas (Technelysium Pty. Ltd.) or 4Peaks (Mekentosj).

2.3. Protein electrophoresis and Western blot analysis

2.3.1. Sodium dodecyl sulphate polyacrylamide gel electrophoresis

Protein samples to be resolved by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) were prepared in Laemmli sample buffer (Laemmli, 1970) (final concentrations: 2 % (w/v) SDS, 10 % (v/v) beta-mercaptoethanol, 20 % (v/v) glycerol, 0.002 % (w/v) bromophenol blue, 0.125 M Tris pH 6.8). The XCell SureLock Mini-Cell Electrophoresis System (Invitrogen) was used for SDS-PAGE. Tris-glycine gels were prepared in either 1 or 1.5 mm Novex Gel Cassettes (Invitrogen) as described by

Sambrook and Russell (2001); a resolving gel consisted of 8-12 % acrylamide/bisacylamide 30 % (37.5:1) mix (National Diagnostics), 375 mM Tris pH 8.8, 0.1 % (w/v) SDS, 0.1 % (w/v) ammonium persulphate (APS), 0.04 % (v/v) TEMED. Isopropanol and water were placed on top of the resolving mix when setting, and then removed and washed with water before addition of the stacking gel mix (5 % acylamide mix, 125 mM Tris pH 6.7, 0.1 % (w/v) SDS, 0.1 % (w/v) APS, 0.1 % (v/v) TEMED) and comb. Wells were washed with water and SDS running buffer (25 mM Tris, 250 mM glycine, 0.01 % (w/v) SDS) and samples loaded along with a molecular weight marker, Precision Plus (BioRad) or Precision Plus Dual Colour (BioRad) for Coomassie staining or Western blot analysis respectively. Electrophoresis was then performed at 150 V in SDS Running Buffer until the dye was running off the bottom of the gel.

2.3.2. Coomassie blue staining

Gels were stained with Coomassie blue solution (50 % (v/v) methanol, 10 % (v/v) acetic acid, 0.5 % (w/v) Coomassie blue) for ~10 minutes with rocking. Coomassie blue solution was removed and the gel washed briefly with water before addition of destaining solution (10 -20 % methanol, 10 % acetic acid). Gels were destained for 3-16 hours until sufficient decoloration, before removing the solution and replacing it with water. For detection of nanogram levels of protein the Colloidal Blue Staining Kit (Invitrogen) was used according to the manufacturer's instructions. Gels were dried using the GelAir Drying System (Biorad) following the manufacturer's instructions.

2.3.3. Western blot analysis

All primary and secondary antibodies used in Western blot analysis are detailed in Tables 2.1 and 2.2 respectively. Samples to be analysed by Western blot were first resolved by SDS-PAGE (section 2.3.2). The electrophoresed gel was washed with water and equilibrated in Transfer buffer (20 mM Tris, 50 mM glycine, 10 % (v/v) methanol) before transferring to polyvinyladine fluoride (PVDF) membrane (Millipore) previously activated with methanol. Transfer was performed using the XCell II Blot Module (Invitrogen) according to the manufacturer's instructions. Proteins were transferred in transfer buffer at 25 V for 60-90 minutes, or 10 V for 10 hours. The membrane was then blocked with Tris buffered saline (TBS: 280 mM NaCl, 20 mM Tris) containing 0.05 % (v/v) Tween 20 and 5 % (w/v) non-fat dried milk (Marvel) for at least 1 hour on a room temperature rocker. The primary antibody was then diluted in fresh blocking buffer and added to the membrane, typically 3 ml in a 50 ml Falcon tube containing the membrane, and incubated at 4 °C overnight on a roller. The membrane was then washed 3 times in TBS supplemented with 0.05 % (w/v) Tween 20 whilst the primary antibody and milk mix

Table 2.1. Primary antibodies

	Antibody	Species	Source	Dilution	Notes
Human	PrimPol	Rabbit	Luis Blanco, CBMSO, Madrid	WB 1:1000, IF 1:200	Made from insoluble prep (*E. coli*)
	PrimPol	Rabbit	Doherty Lab, GDSC, Brighton	WB 1:1000, IF 1:200	Made from soluble prep (Insect cells)
	PolG1	Rabbit	Abcam (ab2969)	WB 1:250	
	PolG2	Rabbit	Ian Holt, MRC, Cambridge	WB 1:1000	
	mtSSB	Rabbit	Valeria Tiranti, Milan, Italy	WB 1:750	
	ATAD3	Rabbit	Ian Holt, MRC, Cambridge	WB 1:40000	
	PCNA (PC10)	Mouse	CRUK	WB 1:1000	
	RPA1 (RPA70-9)	Mouse	Calbiochem (NA13)	WB 1:200	
	RPA2 (RPA31-19)	Mouse	Calbiochem (NA18)	WB 1:500, IF 1:200	
	Chk1 (2G1D5)	Mouse	Cell Signalling (2360)	WB 1:1000	
	p-Chk1 Ser345 (133D3)	Rabbit	Cell Signalling (2348)	WB 1:1000	
	Histone H1.2	Rabbit	Abcam (ab17677)	WB 1:1000	
	β-Actin (AC74)	Mouse	Sigma (A5316)	WB 1:7500	
	α-Tubulin (B-5-1-2)	Mouse	Sigma (T5168)	WB 1:3000	
T. brucei	Rad51	Rabbit	Horn Lab, LSHTM, London	IF 1:500	Proudfoot and McCulloch, 2005
Tags	HA (HA.C5)	Mouse	Abcam (ab18181)	WB 1:1000 IF 1:200	
	FLAG (M2)	Mouse	Sigma (F3165	WB 1:750	
	6-His HRP	-	Abcam (ab1187)	WB 1:5000	
	c-Myc (9E10)	Mouse	Santa Cruz (sc-40)	WB 1:1000 IF 1:200	

Primary antibodies used in this thesis for Western blot (WB) and immunofluoresence (IF) analysis against human proteins, *T. brucei* proteins, and epitope tags are shown. The two anti-PrimPol antibodies were used interchangeably during this thesis. In brackets next to the name is the clone number and next to the supplier the catalogue number for commercial antibodies. Anti-6-His antibody is conjugated to horseradish peroxidase (HRP) and therefore does not require a secondary antibody. CBMSO, Centro de Biología Molecular Severo Ochoa; GDSC, Genome Damage and Stability Centre; MRC, Medical Research Council; LSHTM, London School of Hygiene and Tropical Medicine.

Table 2.2. Secondary antibodies

Antibody	Source	Dilution
Rabbit HRP	Abcam (ab6721)	WB 1:5000
Mouse HRP	Abcam (ab6728)	WB 1:5000
Rabbit Alexa Fluor 488 (green)	Invitrogen	IF 1:2000
Rabbit Alexa Fluor 594 (red)	Invitrogen	IF 1:2000
Mouse Alexa Fluor 488 (green)	Invitrogen	IF 1:2000
Mouse Alexa Fluor 594 (red)	Invitrogen	IF 1:2000

Secondary antibodies used in this thesis for Western blot (WB) and immunofluoresence (IF) analysis. HRP, horseradish peroxidase; Alexa Fluor 488, fluorescent dye of absorbance/emission maxima of 495/519 nm visible as green fluoresence; Alexa Fluor 595, fluorescent dye with absorbance/emission maxima of 590/617 nm visible as red fluoresence.

could be stored at -20 ºC for further use. Following washes, the secondary antibody which is conjugated to horse radish peroxidase (HRP) was added in blocking buffer and incubated on a room temperature rocker for ~1 hour. The membrane was then washed a further 3 times before chemiluminescent detection with Amersham ECL Western Blotting Detection Reagent (GE Healthcare) was performed according to the manufacturer's instructions. Light emission was captured with autoradiography film (GE Healthcare) using a Xograph compact 4 automatic X-Ray film processor.

2.3.4. Far-Western blot analysis of slot-blotted recombinant proteins

This method was used to test for protein-protein interactions *in vitro*. A recombinant protein was first slot-blotted using a slot blot manifold (GE Healthcare) onto a methanol-activated PVDF membrane according to the manufacturer's instructions. Typically, increasing concentrations of recombinant protein (between 250 ng and 2 µg) were blotted onto the membrane. For far-Western analysis the membrane was first blocked as described previously for Western blot analysis (section 2.3.3), and then incubated with blocking buffer containing 0.1 µg/ml (typically 10 ml was used) of the candidate interacting recombinant protein. The membrane was then washed, probed with primary and secondary antibodies, and subjected to chemiluminescent detection as described previously (section 2.3.3). The primary antibody used would be specific for the candidate interacting recombinant protein, and chemiluminescence detected if interaction between the two proteins occurred.

2.4. *T. brucei* methods

All plasmid constructs and the primers used to generate stable *T. brucei* strains are detailed in Tables 2.3 and 2.4 respectively.

2.4.1. Strains and maintenance

Bloodstream-form *T. brucei*, Lister 427, MiTat 1.2, clone 221a, and the transgenic derivative 2T1-strain (Alsford and Horn, 2008) were maintained in Hirumi's modified Iscove's (HMI)-11 medium (Hirumi and Hirumi, 1994), which is the same as HMI-9 but lacks Serum plus. Cells were typically maintained in 36 ml medium in a 25 cm^2 flask. Cells were split every day to a cell density of 1 x 10^5 cells/ml (counting on a haemocytometer) in fresh pre-warmed HMI-11 medium and incubated at 37 ºC with 5 % CO_2. For storage, 0.9 ml of culture (~1 x 10^6 cells/ml) was added to 100 µl of 100 % glycerol (giving 10 % final) in a cryotube, cooled to -80ºC, and then placed in liquid nitrogen.

Table 2.3. Plasmid constructs for *T. brucei*

Construct (vector:insert)	Source	Purpose	Construction	Notes
pRPa^{iSL}	David Horn (LSHTM, London)	Generate Tet-inducible RNAi strain		Alsford and Horn, 2008
pRPa^{iSL}:TbPrimPol1	This thesis	TbPrimPol1 RNAi	MCS1, KpnI - BamHi; MCS2, XbaI - XhoI	
pRPa^{iSL}:TbPrimPol2	This thesis	TbPrimPol2 RNAi	MCS1, KpnI - BamHi; MCS2, XbaI - XhoI	
pRPa^{iSL}:TbPriL	This thesis	TbPriL RNAi	MCS1, KpnI - BamHi; MCS2, XbaI - XhoI	
pNAT^{x12M}	David Horn (LSHTM, London)	*In situ* 12xMyc-tag to native allele		Alsford and Horn, 2008
pNAT^{x12M}:TbPrimPol1	This thesis	TbPrimPol1 tag	AscI - XbaI	
pNAT^{x12M}:TbPrimPol2	This thesis	TbPrimPol2 tag	HindIII - XbaI	
pNAT^{x12M}:TbPriL	This thesis	TbPriL tag	HindIII - XbaI	
pRPa^{i6M}	David Horn (LSHTM, London)	Inducible expression of 6xMyc tagged protein		
pRPa^{i6M}:TbPrimPol2	This thesis	TbPrimPol2-6Myc inducible expression	PacI - XbaI/AvrII	Tag C-terminal
pPAC BS-KS	Richard McCulloch (WTCMP, Glasgow)	Delete targetted allele with puromycin cassette		
pPAC BS-KS:TbPrimPol2	This thesis	Delete TbPrimPol2 allele	5', SacI - XbaI; 3', ApaI - KpnI	

Constructs used and generated in this thesis for the *T.* brucei study. RP, *RRNA* promoter controls the expression cassette; NAT, native locus controls expression; a, contains AscI sites for linearisation; i, inducible; SL, stem-loop; M, c-Myc epitope; BS, bluescript; KS, orientation of polylinker – Kpn1 nearest and SacI furthest from LacZ promoter; PAC, puromycin resistance cassette. TbPrimPol1, TbPrimPol2, and TbPriL, are *T. brucei* genes Tb927.5.4070, Tb927.10.2520, and Tb927.10.3110, respectively. LSHTM, London School of Hygiene and Tropical Medicine; WTCMP, Wellcome Trust Centre for Molecular Parasitology. Restriction sites in 'Construction' column refer to the sites that the gene of interest, or fragment of gene of interest, is cloned between. MCS, multiple cloning site; 5', 5' target; 3', 3' target.

Table 2.4. Primers used to generate plasmid constructs for *T. brucei*

Name	Sequence 5'-3'	Restriction sites	Construct (vector:insert)	Notes
P1RiF	GATC**TCTAGA**GGATCCAGACGGCATAATCTGCTCA	**XbaI** BamHI	pRPa^{iSL}:TbPrimPol2	
P1RiR	GATC**CTCGAG**GTACCATGGCTGTTACTTTGGGCAC	**XhoI** KpnI		
P1CF	GATCAAGCTTCATCATGTGGAGTGATT	HindIII	pNAT^{x12M}:TbPrimPol2	
P1CR	GATCTCTAGACCACTGTTGGACGATATCT	XbaI		
P1XCF	GATCTTAATTAATGGGTCCGAAAACTCAGG	PacI	pRPa^{i6M}:TbPrimPol2	
P1XCR	GATCCCTAGGCCACTGTTGGACGATATCT	AvrII		
5160F5	GATCGAGCTCGGCTGCTATGCATCTGTCA	SacI	pPAC BS-KS:TbPrimPol2	Cloning of target from 5' UTR of TbPrimPol2 gene
5160R5	GATCTCTAGATATTCGTATAGGCGGGTTG	XbaI		
5160F3	GATCGGGCCCAGGTGGTTTAGGTTTGTATTG	ApaI		Cloning of target from 3' UTR of TbPrimPol2 gene
5160R3	GATCGGTACCGCAGCCTGGAGGTCCAAG	KpnI		
P2RiF	GATC**TCTAGA**GGATCCTGCATCGCTCATTACTTTGC	**XbaI** BamHI	pRPa^{iSL}:TbPrimPol1	
P2RiR	GATC**CTCGAG**GTACCGTTGCGGTGGTCCTCTGTAT	**XhoI** KpnI		
P2CF	GATCGGCGCGCCGCTTGGTTTGGGCG	AscI	pNAT^{x12M}:TbPrimPol1	
P2CR	GATCTCTAGAAGCGTCCTGCGACGGAC	XbaI		
P3RiF	GATC**TCTAGA**GGATCCTTTTGTTTTCTGCATGAGCG	**XbaI** BamHI	pRPa^{iSL}:TbPriL	
P3RiR	GATC**CTCGAG**GTACCAAGTGTGTCTGCGCAAGATG	**XhoI** KpnI		
P3CF	GATCAAGCTTCGGCTGGTGCTG	HindIII	pNAT^{x12M}:TbPriL	
P3CR	GATCTCTAGACAAAGAATCACGCGGACGG	XbaI		

Primers designed and used in this thesis to generate constructs for *T. brucei s*tudy. Sequences of the endonuclease restriction sites contained in the primers are indicated by underlining or are in bold. 'Construct' column refers to the construct generated with the primer sets. pRPa^{iSL} is for the generation of tetracycline-inducible RNAi *T.* brucei strains; pNAT^{x12M} is to add 12 c-Myc epitopes to a native allele; pRPa^{i6M} is to generate a strain with tetracycline-inducible over-expression of 6 cMyc epitope fusion protein; pPAC BS-KS is to delete an allele. TbPrimPol1, TbPrimPol2, and TbPriL, are *T. brucei* genes Tb927.5.4070, Tb927.10.2520, and Tb927.10.3110, respectively. UTR, un-translated region.

2.4.2. Stable transformation

Stable transformation of *T. brucei* was performed as described by Alsford and Horn (2008). The constructs to be transfected were prepared by MidiPrep (Qiagen) (section 2.1.3), linearised by restriction endonuclease digest (section 2.2.3), and then subjected to phenol-chloroform extraction. Approximately 2.5 x 10^7 cells were collected by centrifugation (1000 g, 10 minutes) and resuspended in ~25 ml cytomix (120 mM KCl, 0.15 mM $CaCl_2$, 10 mM K_2HPO_4/KH_2PO_4 pH 7.6, 25 mM HEPES pH 7.6, 2 mM EGTA pH 7.6, 5 mM $MgCl_2$, 0.5 % glucose, 100 µg/ml BSA, 1 mM hypoxanthine) (adapted from van der Hoff *et al*, 1992). Cells were then centrifugated again and resuspended to ~6 x 10^7 cells/ml in cytomix. Transfections were performed in duplicate, 0.4 ml of cell suspension was added to ~3 µg DNA which was mixed and electroporated (1.4 kV, 25 µF; Gene Pulser II, BioRad) before addition of 36 ml HMI-11 medium. After 6 hours, selection antibiotics were added and cells divided into 12/24 well plates and clones allowed to grow over ~5 days. Antibiotic concentrations were as follows: blasticidine 10 µg/ml, hygromycin 2.5 µg/ml, phleomycin 2 µg/ml, puromycin 2 µg/ml. Tetracycline was used for induction of dsRNA or recombinant protein expression at a concentration of 1 µg/ml.

2.4.3. Preparation of genomic DNA

Genomic DNA was prepared from *T. brucei* using the DNeasy Blood and Tissue Kit (Qiagen). Cells from 10 ml of a mid-log phase culture (~1 x 10^6 cells/ml) were collected by centrifugation (1000g, 10 minutes) and washed in PBS before extraction of genomic DNA following the manufacturer's protocol.

2.4.4. Preparation of cell lysates

Cells from 10 ml of a mid-log phase culture (~1 x 10^6 cells/ml) were collected by centrifugation (1000 g, 10 minutes), washed in PBS, and resuspended to a final concentration of 1 x 10^7 cells per 100 µl of Laemmli sample buffer. Lysates were incubated at 95 ºC for 5 minutes and 10 µl resolved by SDS-PAGE (section 2.3.1), the remaining sample was stored at -20 ºC.

2.4.5. Immunofluorescence microscopy

Cells from 1 ml of mid-log phase culture (~1 x 10^6 cells/ml) were collected by centrifugation (6000 g, 1 minute), washed in 1 ml PBS, and then resuspended gently in the remaining 50 µl PBS. To this, 50 µl of 4 % (v/v) formaldehyde in PBS was added to fix the cells in a final concentration of 2 % (v/v) formaldehyde, and was then stored at 4

ºC for 1 hour to 7 days. The cells were washed twice with 1 ml ice cold PBS, once with 1 % BSA in PBS, and finally resuspended in the remaining ~30 µl 1 % BSA. Around 5 µl was transferred to a well of a 12-well slide (MP Biomedicals) and allowed to dry overnight.

Primary and secondary antibodies used in immunofluorescence analysis are detailed in Tables 2.1 and 2.2 respectively. For immunofluorescent staining the slides were first washed in PBS for 5 minutes in Coplin jars. Cells were permeabilised with 0.5 % Triton X-100 in PBS for 15 minutes then washed with PBS. Excess PBS was removed from the slide and to the wells 50 % (v/v) foetal bovine serum (FBS) in PBS was added for at least 15 minutes to block non-specific binding of antibodies. The FBS was removed and the primary antibody diluted in 3 % (v/v) FBS in PBS was added and left for at least 45 minutes. Following 2 washes with PBS in the Coplin jar, excess PBS was removed and secondary antibody diluted in 3 % (v/v) FBS in PBS was added to the wells of the slide and incubated in the dark for 45 minutes, before washing as previous. The cells were counterstained with DAPI (final concentration 0.1 µg/ml) in Vectashield (Vector Laboratories Inc.) and a cover slip placed over the wells and sealed with nail varnish. Images were captured using a Nikon Eclipse E600 epifluorescence microscope (Nikon) in conjunction with a Coolsnap FX (Photometrics) charge-coupled device (CCD) camera and processed in Metamorph 5.0 (Photometrics).

2.4.6. TUNEL staining

Terminal deoxynucleotidyl transferase dUTP nick end labelling (TUNEL) was performed to detect DNA strand breaks in *T. brucei*, using the Fluorescein *In Situ* Cell Death Detection Kit (Roche). Cells were fixed and mounted on microscope slides as described for immunofluorescence microscopy (section 2.4.5), and the TUNEL staining was performed according to the manufacturer's instructions. Images were captured as described previously (section 2.4.5).

2.4.7. Flow cytometry

Cells from 1 ml of mid-log phase culture (~1 x 10^6 cells/ml) were collected by centrifugation (6000 g, 1 minute), washed in cold PBS, and resuspended in the remaining 300 µl PBS. To this, 700 µl ice cold 100 % methanol was added (therefore fixing in 70 % methanol) and mixed by gently inversion before storing overnight at 4 ºC. Fixed cells were collected by centrifugation (200 g, 10 minutes, 4 ºC), washed in PBS, and resuspended in 1 ml PBS before transfer to a FACS tube (BD Biosciences) and addition of 10 µg/ml RNase A and 10 µg/ml propidium iodide (PI). The tube was shaken gently to

mix and incubated in the dark for 45 minutes at 37 ºC. A FACScalibur (BD Biosciences) was used to analyse samples with CellQuest software and detector FL2-A with an Amp gain value of 1.75. Data analysis was carried out using FlowJo (Treestar).

2.5. Preparation of recombinant proteins in *E. coli*

Plasmid constructs used and generated in this thesis for the production of recombinant proteins in *E. coli*, and the primers used to make these constructs, are detailed in chapters 2.5 and 2.6 respectively

2.5.1. Preparation of chemically competent B834s

The *E. coli* strain B834 (DE3) pLysS (B834s) (Novagen) was used to over-express recombinant proteins used in this thesis. The TSS method (Chung *et al*, 1989) was used to produce competent cells, which requires no heat shock during transformation, an advantage given that cells containing pLysS are more susceptible to lysis. A fresh overnight culture of B834s was diluted 1:100 into LB medium and incubated at 37 ºC with shaking until OD_{600} of 0.3-0.4. An equal volume of 2x transformation and storage solution (2x TSS; LB medium with 20 % PEG 3350, 10 % DMSO, 100 mM $MgSO_4$, pH 6.5) was added, mixed gently, and the mixture aliquoted and snap frozen before storage at -80ºC.

2.5.2. Transformation of competent B843s

A 50 µl aliquot of competent B834s was thawed on ice and 1 µl of MiniPrep plasmid DNA (section 2.1.3) of the appropriate expression construct was added. Following a 15 minutes incubation on ice, 1 ml LB medium was added and the cell DNA mixture incubated at 37 ºC with shaking for at least 40 minutes. The cells were collected by centrifugation, resuspended in ~100 µl LB medium, and plated on LB agar plates containing 30 µg/ml chloramphenicol (for pLysS of B834s) and the antibiotic resistance of the expression construct. Colonies were allowed to grow overnight at 37 ºC. Plates containing transformants were stored at 4 ºC or for long term storage glycerol stocks were made. Briefly, 750 µl overnight culture was added to 250 µl sterile 60% % glycerol in a cyrotube and stored at -80 ºC.

2.5.3. Recombinant protein expression trials

B834s transformed with the appropriate expression construct (section 2.5.2), either from a fresh single colony or a glycerol stock, was used to inoculate 50 ml of LB medium containing 30 µg/ml chloramphenicol and the antibiotic resistance present on the

Table 2.5. Plasmid constructs for protein expression in *E. coli*

Construct (vector:insert)	Source	Purpose	Construction	Notes
pET28a	Novagen	IPTG inducible expression in *E. coli*		
pET28a:TbPrimPol1	This thesis	6xHis-TbPrimPol1 expression	NdeI - BamHI	Tag N-terminal
pET28a:TbPrimPol2	This thesis	6xHis-TbPrimPol2 expression	NdeI - EcoRI	Tag N-terminal
pET28a:TbPrimPol1-AxA	This thesis	6xHis-TbPrimPol1-AxA expression	SDM - D165A E167A	Tag N-terminal
pET28a:TbPrimPol2-AxA	This thesis	6xHis-TbPrimPol2-AxA expression	SDM - D193 D195A	Tag N-terminal
pRUN:POLG2	Ian Holt (MRC, Cambridge)	10xHis-PolG2 expression in *E. coli*		Tag N-terminal

Constructs used and generated in this thesis for the expression of recombinant proteins in *E. coli*. Both pET28a and pRUN contain a T7 promoter with a *lac* operon, allowing IPTG-inducible expression of the downstream gene. 'Construction' column either contains endonuclease restriction sites that the gene of interest is cloned between, or point-mutations introduced by site-directed mutagenesis (SDM). TbPrimPol1, TbPrimPol2, and TbPriL, are *T. brucei* genes Tb927.5.4070, Tb927.10.2520, and Tb927.10.3110, respectively. POLG2 is human gene Polymerase (DNA directed) gamma 2 accessory subunit.

Table 2.6. Primers used to generate plasmid constructs for protein expression in *E. coli*

Name	Sequence 5'-3'	Restriction sites	Construct (vector:insert)	Notes
Tb_5.4070_F	gctCATATGctctggtgtagggcaacg	NdeI	pET28a:TbPrimPol1	
Tb_5.4070_R	cGGATCCctaagcgtcctgcgacg	BamHI		
Tb_70.5160_F	cGGATCCctaagcgtcctgcgacg	BamHI	pET28a:TbPrimPol2	
Tb_70.5160_R	cGAATTCtcaccactgttggacgatatc	EcoRI		
Tb4070_AXA_F	GTATTTGG**cc**GTTG**ct**CGGGAGAGG		pET28a:TbPrimPol1-AxA	
Tb4070_AXA_R	GTATTTGG**cc**GTTG**ct**CGGGAGAGG			
Tb5160_AXA_SJ_F	CCTTTTTTCG**cg**ATAG**c**CTGCAGCCCACCGTTTGAGTGGTTGAATTTTCC		pET28a:TbPrimPol2-AxA	Creates NruI restriction site for screening
Tb5160_AXA_SJ_R	GTGGGCTGCAG**g**CTAT**cg**CGAAAAAAGGGTCAACTGCGGCGTCTTTATC			

Primers designed and used in this thesis to generate plasmid constructs for expression of recombinant proteins in *E. coli*. Sequences of the endonuclease restriction sites contained in the primers are underlined, and highlighted in bold are the mutated bases for site-directed mutagenesis PCR. 'Construct' column refers to the construct generated with the primer sets. pET28a is for the over-expression of proteins in *E. coli*. TbPrimPol1 and TbPrimPol2 are *T. brucei* genes Tb927.5.4070 and Tb927.10.2520 respectively. TbPrimPol1-AxA refers to mutations D165A E167A, and TbPrimPol2-AxA refers to mutations D193 D195A.

expression construct. Following an overnight incubation at 37 ºC with shaking, the culture was diluted 1:100 into fresh LB medium (~250 ml) containing the relevant antibiotics and grown at 37 ºC until an OD_{600} of 0.6. The culture was then cooled on ice, a pre-induction sample taken (~1ml of cells spun down), and then divided between multiple flasks depending on induction conditions to be tested. Typically short inductions at 30 and 37 ºC were tried, longer inductions were performed at 25 ºC, and overnight inductions at 18 ºC. If desired, at different time-points (every hour), the OD_{600} was measured, 1 ml of cells spun down and lysed in Laemmli sample buffer for whole cell extract, and a further 10 ml culture centrifugated for fractionation into soluble and insoluble samples (section 2.5.4). Expression was first tested on whole cell lysate, loading an equivalent to 0.15 OD_{600} equalling ~40 µg, and if expression visible, the proportion of soluble protein could be determined by processing the 10 ml pellet. Samples were analysed by SDS-PAGE and Coomassie staining, as well as detection of the epitope tag by Western blot (section 2.3).

2.5.4. Protein expression and preparation of *E. coli* lysate

For purification of recombinant proteins, first an overnight culture of B834s transformed with the appropriate expression construct was diluted 1:100 into fresh LB medium, typically 3 L divided into six 0.5 L cultures, containing the relevant antibiotics. The cultures were grown at 37 ºC until the OD_{600} was 0.6-0.8 (~3-4 hours), then cooled rapidly on ice for 10 minutes before addition of Isopropyl β-D-1-thiogalactopyranoside (IPTG) and returned to the incubator for the desired amount of time. Cells were collected by centrifugation (5000 rpm, 10 minutes, 4 ºC) and lysed into a small volume of buffer A (according to the chromatography column to be used) (10 ml per litre of culture) and stored at -80 ºC.

The cell lysate was thawed on ice and further diluted with buffer A, up to 30 ml per litre of culture grown, then placed on a stirrer on ice for ~30 minutes with the addition of 0.1 mg/ml lysozyme. The lysate was then sonicated (Vibra-Cell Sonicator) on ice, 4 rounds of 20 second pulses at 30 % amplitude, resting for at least 1 minute in between pulses, before being cleared by centrifugation (18 000 rpm, 1 hour, 4 ºC; Sorvall RC26 Plus, SS-34 rotor). The soluble fraction was then passed through a 0.45 µm filter (Millipore) and kept on ice for subsequent chromatography analyses.

2.5.5. Immobilised metal affinity chromatography

Immobilised metal affinity chromatography (IMAC) was used to purify proteins fused to a histidine tag. Chromatographic columns were run using the ÄKTAprime system (GE

Healthcare), and all buffers and cell lysate were kept on ice. A column charged with 5 ml of nickel-nitrilotriacetic acid (Ni^{2+}-NTA) agarose resin (Qiagen) was equilibrated in IMAC buffer A (10 % (v/v) glycerol, 0.5 M NaCl, 50 mM Tris pH 7.5, 5 mM β-mercaptoethanol, 30 mM imidazole, 17 µg/ml PMSF, 34 µg/ml benzamidine) and then the soluble *E. coli* lysate was loaded onto the column at a flow rate of 2-3 ml/min. The column was washed extensively at 5 ml/min with IMAC buffer A until all the unbound proteins were removed, which was determined by the absorbance at 280 nm (A_{280}) returning to background levels. The column was then washed with 10 % IMAC buffer B (same as buffer A except addition of 300 mM imidazole) and then bound proteins eluted at 2-3 ml/min with 100 % IMAC buffer B (300 mM imidazole). Samples were analysed by SDS-PAGE with Coomassie staining and anti-His Western blot analysis (section 2.3).

2.5.6. Heparin affinity chromatography

Heparin agarose is commonly used to purify DNA binding proteins as it is a phosphate mimetic. If the IMAC elution was to be further purified by heparin chromatography the elution was first diluted at least 10-fold to reduce the salt concentration, as this would prevent binding. The elution was diluted in heparin buffer A (10 % (v/v) glycerol, 50 mM Tris pH 7.5, 2 mM DTT) and loaded onto a 5 ml Hi-Trap Heparin sepharose column (GE Healthcare) which was pre-equilibrated in buffer A. Following an extensive wash with buffer A, and 10 % buffer B (same as A except 1 M NaCl), step elution was performed.

2.5.7. Storage of recombinant proteins

Purified recombinant protein was concentrated using a Vivaspin sample concentrator (GE Healthcare) with the appropriate molecular weight filter and the concentration measured at A_{280} using a Nanodrop spectrophotometer (Thermo Scientific). Once the desired concentration reached, the purified proteins were aliquoted in small volumes (for single use) and snap frozen in liquid nitrogen, then stored at -80 ^{o}C.

2.6. Primer extension assay

The primer extension assay and associated protocols were modified from Jozwiakowski and Connolly, 2011.

2.6.1. Primer-template substrates

The DNA oligomers used to prepare the synthetic primer-template substrates were designed by Stanislaw Jozwiakowski (Sussex, UK) and were HPLC grade manufactured by ATDbio. The DNA oligomer containing the thymine-thymine pyrimidine (6-4)

pyrimidone photoproduct was a gift from Alan Lehmann (Sussex, UK), and the 3-deaza-3-methyl-2'-deoxyadenosine (3dMeA) containing oligomer was a gift from Roger Woodgate (NIH, Maryland, USA) (Plosky *et al*, 2008). All primers contain a 5' hexachlorofluorescein label and the sequences of all the oligonucleotides used in the primer extension assays in this thesis are detailed in Table 2.7.

2.6.2. Annealing of primer-template substrates

To anneal primer-template substrates the labelled DNA primer and the DNA template were mixed in an equimolar ratio in annealing buffer (50 mM Tris pH 8, 50 mM NaCl, 1 mM EDTA) and incubated at 95 ºC for 3 minutes before allowing to cool down to room temperature for a further 30 minutes. Annealed substrates (200 nM) were stored at -20 ºC.

2.6.3. Primer extension assay

Reactions were assembled at room temperature; to 1x reaction buffer (NEBuffer 1, New England Biolabs; 10 mM Bis-Tris-Propane pH 7, 10 mM $MgCl_2$, 1 mM DTT) 20 nM primer-template substrate was added with 200 µM dNTP(s) (Roche) and lastly 100 - 125 nM polymerase. Volume for a single reaction was 20 µl, which was scaled up if multiple time-points were performed. Reactions were incubated at 37 ºC for the desired time (10-30 minutes) and then quenched with an equal volume of 2x stop buffer (95 % formamide, 40 mM EDTA, 200 nM competitor oligonucleotide, with bromophenol blue and xylene cyanol), either by adding directly to the tube or, if multiple time-points were being performed, 20 µl removed and added to a clean tube containing 20 µl 2x stop buffer. The competitor oligonucleotide in the stop buffer is complementary to the template strand in the assay, and this prevents re-annealing of the labelled reaction products with the template strand, which would interfere with gel electrophoresis analysis. Quenched reactions were then heated to 95 ºC for 3 minutes and allowed to cool down before half of the reaction was resolved by DNA-PAGE, and the remainder stored at -20 ºC.

2.6.4. DNA-polyacrylamide gel electrophoresis

The products of primer extension reactions were resolved by denaturing urea-PAGE. Gels consisted of 15 % polyacrylamide bis-acrylamide (19:1) solution (National Diagnostics), 7 M urea, TBE (89 mM Tris, 2 mM EDTA, 0.89 M Boric Acid, pH 8.3), 0.1 % APS, 0.03 % TEMED. Gels were cast between 16.5 x 28 cm glass plates with 0.75 mm spacers and a 20-well comb, and resolved on a Dual Vertical Slab Gel Kit (CBS Scientific). Gels were pre-run at a constant wattage of 15 watts for 30 minutes in TBE, before loading the samples and electrophoresis for 3-4 hours at 15-17 watts. To detect

Table 2.7. Oligonucleotides used in this thesis for the primer extension assays

	Length	Sequence (5'-3')	Lesion
Template	50	CGCGCAGGGCGCACAACAGCCTTGAAGACCGAACGACCGAACAGACGACA	No damage
	50	CGCGCAGGGCGCACAACAGCCTTGAAGACCGAACGACCGAACAGACGACA	CDP (TT)
	50	CGCGCAGGGCGCACAACAGCCGTGAAGACCGAACGACCGAACAGACGACA	8-oxo-G (G)
	50	CGCGCAGGGCGCACAACAGCC_TGAAGACCGAACGACCGAACAGACGACA	AP site (_)
Competitor	50	TGTCGTCTGTTCGGTCGTTCGGTCTTCAAGGCTGTTGTGCGCCCTGCGCG	
Primer (5' HEX label)	20	TGTCGTCTGTTCGGTCGTTC	
	28	TGTCGTCTGTTCGGTCGTTCGGTCTTCA	
	29	TGTCGTCTGTTCGGTCGTTCGGTCTTCAA	
	29	TGTCGTCTGTTCGGTCGTTCGGTCTTCAC	
Template	30	CTCGTCAGCATCTTCATCATACAGTCAGTG	No damage
	30	CTCGTCAGCATCTTCATCATACAGTCAGTG	6-4 PP (TT)
Competitor	30	CACTGACTGTATGATGAAGATGCTGACGAG	
Primer (5' HEX label)	16	CACTGACTGTATGATG	
	17	CACTGACTGTATGATGT	
Template	18	AGACGACGTCCTGTAGCC	No damage
	18	AGACGACGTCCTGTAGCC	3dMeA (A)
Competitor	18	GGCTACAGGACGTCGTCT	
Primer (5' HEX label)	12	GGCTACAGGACG	

Primer, template, and competitor oligonucleotides, shown in their compatible groups, used in the primer extension assays in this thesis. DNA lesions in templates are indicated in red text: CPD, cyclobutane pyrimidine dimer; 8-oxo-G, 8-oxo-guanine; AP site; abasic site; 6-4PP, pyrimidine-pyrimidone 6-4 photoproduct; 3dMeA, 3-deaza 3-methyladenine, a stable anaologue of 3-methyl adenine. Competitor oligonucleotides are complementary to the template strand and are added when the reaction is quenched to prevent re-annealing of labelled products with the template, which interferes with gel electrophoresis analysis. Primers contain a 5’ hexachlorofluorescein (HEX) label, which has a peak emission at 555 nm.

the hexachlorofluorescein label of the reaction products, the gel was scanned (Cy3, 532 nm) using a FLA-1500 scanner (FUJI).

2.7. Human cell culture

The cell media, supplements, and antibiotics, were purchased from Gibco Invitrogen. Cell culture flasks, plates, dishes, and cryotubes were purchased from Nunc. Plasmid constructs used and generated in this thesis in cultured human cells, and the primers used to make these constructs, are detailed in Tables 2.8 and 2.9 respectively.

2.7.1. Cell lines

Osteosarcoma U2OS and 143B cells, human embryonic kidney (HEK)-293 cells, and retinal pigment epithelium (RPE) cells were cultured in Dulbecco's modified Eagle Medium (DMEM) supplemented with 10 % (v/v) foetal calf serum (FCS), 100 U/ml penicillin, 100 mg/ml streptomycin, and 2 mM L-glutamine. Flp-In T-REx-293 cells, engineered for inducible expression of for epitope tagged PrimPol, were cultured in the same media containing 15 µg/ml Blasticidine and 100 µg/ml Hygromycin, whilst the parental cells were cultured in 100 µg/ml Zeocin and 15µg/ml Blasticidine. SV40 transformed normal (MRC5V1) and XP-V (XP30RO(sv), (gift from Alan Lehmann, Sussex, UK; Cleaver *et al*, 1999) fibroblasts were cultured in Minimum Essential Media (MEM) supplemented with 15 % FCS and penicillin, streptomycin, and L-glutamine as above. In this thesis these cells are referred to as MRC5 and XP30RO. All cells were grown in a 37 ^{o}C incubator with a humidified atmosphere containing 5 % CO_2.

2.7.2. Cell maintenance

Cells were maintained in a 75 cm^2 cell culture flask with 10 ml medium and split every 3-5 days. At each passage, medium was first removed and the cells washed with phosphate buffered saline (PBS, 137 mM NaCl, 0.3 mM KCl, 10 mM Na_2HPO_4, 1.7 mM KH_2PO_4) before incubation with trypsin (250 mg/ml in PBS) for ~5 minutes at 37 ^{o}C. Pre-warmed medium was then added and cells detached by gentle agitation or pipetting, and recovered by centrifugation and the cell pellet resuspended in fresh medium. An appropriate amount of the cell suspension was transferred to a clean flask and diluted into fresh medium before being returned to the incubator, typically cells were split 1:8 or 1:10. HEK-293 and the Flp-In T-REx-293 cells did not require trypsinisation, and were resuspended in PBS instead.

Table 2.8. Plasmid constructs for cultured human cells

Construct (vector:insert)	Source	Purpose	Construction	Notes
pEGFP-C1	Clontech	Transient expression of N-ter eGFP tagged protein		
pEGFP-C1:PrimPol	Andrew Green, Doherty Lab	eGFP-PrimPol expression		
pEGFP-N1	Clontech	Transient expression of C-ter eGFP tagged protein		
pEGFP-N1:PrimPol	Andrew Green, Doherty Lab	PrimPol-eGFP expression		
pCI-Neo	Promega	Transient expression of protein		
pCI-Neo:PrimPol-HA	Andrew Green, Doherty Lab	PrimPol-HA expression	XhoI - MluI	HA-tag cloned between XbaI - NotI
pcDNA5/FRT/TO	Invitrogen	Generate stable, inducible expression cell line		
pcDNA5/FRT/TO:PrimPol-HA	This thesis	Inducible PrimPol-HA expression	BamHI - XhoI	Tag C-terminal
pcDNA5/FRT/TO:FLAG-Strep-II-Tag	Ian Holt (MRC, Cambridge)			FLAG-*Strep*-II-Tag sequence between XhoI and ApaI
pcDNA5/FRT/TO:PrimPol-FLAG-Strep-II-Tag	This thesis	Inducible PrimPol-FLAG-Strep-Tag-II expression	BamHI - XhoI	Tag C-terminal

Constructs used and generated in this thesis for the over-expression of recombinant proteins in cultured human cells. EGFP, enhanced green fluorescent protein, FRT, Flp Recombinase Target site; TO, Tet-operator. All constructs use the cytomegalovirus (CMV) immediate-early promoter, pCI-Neo and pEGFP allow constitutive expression of proteins, while pcDNA5/FRT/TO contains a allows tetracyline/doxycycline-inducible expression. 'Construction' column contains endonuclease restriction sites that the gene of interest is cloned between. PrimPol refers to human gene coiled-coil domain containing protein (CCDC)111.

Table 2.9. Primers used to generate plasmid constructs for cultured human cells

Name	Sequence 5'-3'	Restriction sites	Construct (vector:insert)	Notes
ccdcHAStbl_F	AGGATCCCTATGAATAGAAAATGGGAAGC	BamHI	pcDNA5/FRT/TO:PrimPol-HA	PCR Template - pCI-Neo:PrimPol-HA
ccdcHAStbl_R	TCTCGAGCTAAGCGTAGTCTGGGACGTCGT	XhoI		
ccdcstrepflagR	CCGCTCGAGCTCTTGTAATACTTCTATAATTAG	XhoI	pcDNA5/FRT/TO:PrimPol-FLAG-Strep-II-Tag	Use with ccdcHAStble_F, sub-clone into pcDNA5/FRT/TO:FLAG-Strep-II-Tag

Primers designed and used in this thesis to generate plasmid constructs for expression of recombinant proteins cultured human cells. Sequences of the endonuclease restriction sites contained in the primers are underlined. 'Construct' column refers to the construct generated with the primer sets. pcDNA5/FRT/TO:PrimPol-HA is to generate a stable cell line with inducible expression of HA-tagged PrimPol; pcDNA5/FRT/TO:PrimPol-FLAG-Strep-II-Tag is to generate a stable cell line with PrimPol fused to tandem FLAG-Strep-Tag-II epitopes. PrimPol refers to human gene coiled-coil domain containing protein (CCDC)111.

2.7.3. Cell storage

To generate cell stocks, cells were collected as described above and the cell pellet resuspended in pre-warmed fresh medium containing 10 % (v/v) DMSO. Usually, a confluent 75 cm^2 flask was divided into four 1 ml aliquots in cryotubes, which were frozen slowly in a cryobomb placed at -80 ºC before being transferred to liquid nitrogen for long term storage. To restart a cell culture, a frozen aliquot of cells was quickly thawed at 37 ºC and added dropwise to 10 ml of pre-warmed fresh medium. To remove the DMSO, the cell suspension was centrifugated and the pellet resuspended in fresh medium, before transferring to a clean 75 cm^2 flask.

2.8. Cultured human cell methods

2.8.1. Preparation of cell lysates

To prepare whole cell extracts, cells were either collected first by centrifugation and then washed with PBS before addition of lysis buffer, or scraped directly into lysis buffer if only a small number of cells were present. Typically, a cell pellet was resuspended in NETN buffer (150 mM NaCl, 5 mM EDTA, 50 mM Tris pH 7.5, 0.5 % NP-40, with Roche protease inhibitor) and incubated on ice for 30 minutes – 1 hour with occasional vortexing, before sonication in a chilled ultrasonic bath and centrifugation (13 000 rpm, 10 minutes, 4 ºC) to remove cell debris. Usually, 50 µl was used for a confluent 3.5 cm dish and 200 µl for a 10 cm dish. To determine the protein concentration the Lowry method (Lowry *et al*, 1951) with the *DC* Protein Assay (BioRad) was used, typically using 5 µl of the lysate and using BSA to generate a standard curve. Usually a 40 µg sample was prepared in Laemmli sample buffer, final volume ~20 µl, and boiled for 5 minutes.

2.8.2. Stable transformation on Flp-In T-REx-293 cells

To generate stable inducible cell lines the Flp-In T-REx system from Invitrogen was used. Parental Flp-In T-REx-293 cells (HEK-293 derivative) and the pcDNA5/FRT/TO and pOG44 vectors were gifts from Ian Holt (MRC, Cambridge). Twenty-four hours prior to transfection cells were counted using a haemocytometer and seeded in a 6-well plate: 2 x 10^5 cells/well giving ~70 % confluency the following day. Co-transfection of pcDNA5/FRT/TO (containing cDNA of interest) and pOG44 in a 1:9 ratio was done with Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. A control mock transfection with water rather than DNA was performed alongside. Forty-eight hours after transfection cells were transferred to 10 cm dishes and the following day antibiotic selection applied (15 µg/ml Blasticidine, 100 µg/ml Hygromycin). Selective medium was replaced every 3-4 days until large colonies had formed and all cells were

dead in the mock transfection (~2 weeks). Dishes were then trypsinised and stocks of the cell lines made. To assess whether the stable transfection was successful cells were seeded in a 6-well plate with media supplemented with 10 ng/ml doxycycline, and 24 hours later, cells were harvested and analysed by Western blot.

2.8.3. RNA interference

Small interfering (si)RNA duplexes targeting PrimPol mRNA were designed by Julie Bianchi (GDSC, Brighton) and manufactured by Invitrogen as Stealth RNAi siRNA. The sequence of the siRNA duplex was 5'-GAGGAAACCGUUGUCCUCAGUGUAU-3', 5'-AUACACUGAGGACAACGGUUUCCUC-3'. Cells were transfected using Lipofectamine RNAiMAX (Invitrogen); 1 µl RNAiMAX was added to a well of a 24-well plate containing 100 µl OPTI-MEM medium (Invitrogen) with 10 nM siRNA duplex (final concentration), to this a 400 µl suspension of 2 x 10^4 cells was added and returned to the incubator (reverse transfection protocol, see manufacturer's guide). Twenty-four hours later a transfection mix composed of 100 µl OPTI-MEM, 1 µl RNAiMAX, and 10 nM siRNA (final) was added to the well before returning to the incubator (forward transfection protocol, see manufacturer's guide). Usually, a well became confluent 48 hours after seeding and initial transfection, and required passaging into an appropriate size dish; experiments were usually performed the following day (72 hours from the starting point). The number of wells seeded was scaled up as required. In some instances only a single round of transfection was performed either when seeding or on the following day, as described above. In this case, 4 x 10^4 cells were seeded and further experiments were performed 48 hours later. Mock transfections that contained water instead of siRNA were performed alongside as a control.

2.8.4. DNA damaging agents and drug treatments

UV-C irradiation at 254 nm was performed with a germicidal lamp at a fluence rate of 0.5 J/m^2/sec. Cells plated in dishes were aspirated of their medium, washed with PBS, and with no liquid present in the dish and lid removed, exposed to the desired dose of UV-C radiation. Pre-warmed media was then added to the cells before returning to the incubator for the appropriate recovery time. Exposure of cells to ionising radiation was performed with 250 kV X-rays at 12 mA with a dose rate of 0.5 Gy/min. Cells plated in dishes were irradiated in their medium with the lid removed. Hydroxyurea (HU) was made fresh prior to use and added to cell medium to a final concentration of 10 mM for 6 hours. Aphidicolin was prepared in DMSO and added to cell medium to a final concentration of 2 µg/ml for 16 hours, and when releasing from block, at least 2 PBS washes were performed before adding pre-warmed fresh medium. Doxycycline, to

induced expression of recombinant PrimPol in Flp-In T-REx-293 cells, was added to a final concentration of 10 ng/ml for 16-24 hours.

2.8.5. UV-C clonogenic survival assay

MRC5 and XP-V fibroblasts were treated with 2 rounds of PrimPol or mock RNAi as detailed previously (section 2.8.3). Six wells of a 24-well plate for each condition (i.e. mock RNAi MRC5, PrimPol RNAi MRC5, etc) were enough to set up the survival and check PrimPol knockdown by Western blot. Forty-eight hours after seeding and initial transfection, cells were trypsinised and counted using haemocytometer and plated on 10 cm dishes. For MRC5: 200, 400, 1000, and 2000, cells were plated; and XP-V: 400, 800, 2000, and 4000, cells plated, all in triplicate, for mock irradiation, 2 J/m^2, 4 J/m^2, and 6 J/m^2 irradiation respectively. Less than 24 hours after plating, cells were exposed to the desired dose of UV-C irradiation and returned to the incubator until large colonies had grown, typically ~10 days. Cells were then washed with PBS, fixed with 100 % ethanol, and washed with distillated water before staining with methylene blue for 10 minutes. Cells were then washed again with water to remove the excess staining, and left to dry before counting using a cell counter and calculating the average percentage survival for each UV-C dose.

2.8.6 Immunofluorescence microscopy

Primary and secondary antibodies used in this thesis for immunofluorescence analysis are detailed in Tables 2.1 and 2.2 respectively. Cells were grown on glass cover slips, which for Flp-In T-REx-293 cells were poly-L-lysine coated (Invitrogen), in typically 3.5 cm dishes with 2-3 cover slips per dish. Cell medium was removed and cells washed in PBS before fixing in 3 % (v/v) paraformadlehyde (in PBS) for 15 minutes. Following fixation, dishes of fixed cells were often stored at 4 ºC in PBS until immnofluorescent staining. Cells were permeabilised with 0.2 % (v/v) Triton X-100 (in PBS) for 10 minutes, then washed with PBS before blocking with 3 % BSA (in PBS) for at least 30 minutes. Primary antibodies were diluted in blocking buffer and added dropwise to the surface of the cover slips, and incubated for at least 1 hour at room temperature. Cells were then washed with PBS before incubation with the secondary antibody also diluted in blocking buffer, for at least 30 minutes in the dark. Following further washes with PBS, the excess PBS was removed from the cover slip before mounting onto a glass slide with Prolong Gold anti-fade (Invitrogen), which contains 4',6-diamidino-2-phenylindole (DAPI). Nail varnish was then used to coat the edge of the cover slips to create a seal and the slides were stored at 4 ºC until analysis. If slides were to be scored, they were analysed on an E400 (Nikon) microscope with a 100x oil immersion objective. For image acquisition,

slides were analysed with a DeltaVison Core (Applied Precision) microscope on either 60x or 100x objective. Images were taken, deconvolved, and processed using softWoRk® (Applied Precision) and transferred to OMERO (Allan *et al*, 2012) for storage and further processing.

To immunofluorescently detect chromatin bound proteins, cells were pre-extracted before fixation with a rapid wash with 0.5 % Triton X-100 in PBS. To visualise mitochondria, cell medium was supplemented with 250 nM MitoTracker Deep Red (Invitrogen) and cells returned to the incubator for 30 minutes before fixation. To detect S-phase cells the Click-iT EdU Alexa Fluor 488 Imaging Kit (Invitrogen) was used according to the manufacturer's instructions. Cell medium was supplemented with 10 μM EdU (final concentration) 20 minutes before fixing.

2.8.7. Flow cytometry

To analyse the DNA content of human cultured cells by flow cytometry, cells were first seeded in a 6-well plate so that they are ~80 % confluent when harvested. Cells were trypsinised and collected by centrifugation, washed with PBS, then resuspended in a small volume of PBS and fixed with ice cold 70% (v/v) ethanol added dropwise with gentle agitation, before storage overnight at -20 ºC. Cells were then washed twice with PBS to remove the ethanol and resuspended in a PBS containing 5 μg/ml propidium iodide, 250 μg/ml RNase and transferred to FACS tubes (BD Biosciences) to be kept overnight in the fridge until analysis. The DNA content of the cells was analysed using a FACSCanto flow cytometer (BD Biosciences) and the DNA content versus cell count was plotted using FACSDiva software (BD Biosciences)

2.8.8. Sub-cellular fractionation

2.8.8.1. Mitochondrial isolation

Mitochondria were isolated from human cells using the Mitochondrial Isolation kit for cultured cells (Thermo Scientific) according to the manufacturer's instructions. The resulting mitochondrial pellet was lysed in RIPA buffer (20 mM Tris pH 8.0, 150 mM NaCl, 1% NP40, 0.5 mM EDTA, 10% glycerol, 0.1% SDS, 1% Na-deoxycholate), and protein concentration was determined as previously described (section 2.8.1). Equivalent concentrations of cytoplasmic fraction and mitochondrial lysate were prepared and analysed by Western blot (section 2.3.3).

2.8.8.2. Protease protection assay

For the protease protection assay, mitochondria were first isolated from cells as previously described (section 2.8.8.1), and the resulting mitochondrial pellet resuspended in 100 µl PBS and divided into 4 eppendorf tubes. An equal volume of protease protection assay buffer (0.6 M mannitol, 10 mM Tris pH 7.4) was added to each tube, and supplemented with either 0.5 % (v/v) Triton X-100, or 0.5 mg/ml Proteinase K (New England Biolabs), or a combination of the two, or neither. After a 30 minute incubation on ice, 10 mM PMSF (final concentration) was added followed immediately by Laemmli sample buffer before boiling the samples. The contents of each tube was then analysed by Western blot (section 2.3.3) with an anti-PolG2 antibody (gift from Ian Holt, MRC, Cambridge) used as a control for a mitochondrial protein.

2.8.8.3. Soluble and insoluble (chromatin) fractionation and DNase treatment

The fractionation protocol was modified from Kannouche *et al*, 2004. Fractionation of Flp-In T-REx-293 cells over-expressing PrimPol required a single confluent 10 cm dish per condition; in order to detect chromatin bound PrimPol, whilst detecting endogenous PrimPol in MRC5 fibroblasts required three 10 cm dishes per condition. Cells were harvested by scraping into 1-2 ml PBS and a quarter of this suspension transferred to an eppendorf tube and pelleted by centrifugation (3000 rpm, 5 minutes, 4 ºC) for the whole cell extract. The cell pellet was resuspended in 50 µl NETN buffer (150 mM NaCl, 5 mM EDTA, 50 mM Tris pH 7.5, 0.5 % NP-40, with Roche protease and phosphatase inhibitor cocktails), incubated on ice before sonication in a chilled ultrasonic bath. The remaining three quarters of the cell suspension was pelleted by centrifugation and resuspended in 150 µl cytoskeletal (CSK) buffer (100 mM NaCl, 300 mM sucrose, 3 mM $MgCl_2$, 10 mM PIPES pH 6.8, 1 mM EGTA, 0.2 % (v/v) Triton X-100, with Roche protease and phosphatase inhibitor cocktails), and incubated on ice for 5 minutes before centrifugation at 13 000 rpm for 10 minutes at 4 ºC. The supernatant, being the soluble fraction, was transferred to a clean eppendorf tube. The insoluble pellet was washed twice with PBS before being resuspended in 150 µl Laemmli sample buffer and boiled for 10 minutes. The protein concentration of the whole cell extract was determined (section 2.8.1) and equivalent amounts of the soluble and insoluble fractions were analysed by Western blot (section 2.3.3). Detection of Histone H1 was used as a control for the fractionation.

For DNase treatment of the Triton X-100 insoluble fraction, the sample was resuspended in 150 µl low salt CSK buffer (same as CSK buffer except 50 mM NaCl) supplemented with 1 µl/ml Benzonase, and incubated at room temperature for 20 minutes with occasional agitation. The remaining insoluble material was then collected

by centrifugation and the soluble fraction (DNA bound) transferred to a clean tube, whilst the insoluble was resuspended in Laemmli sample buffer and boiled. Equivalent amounts to the whole cell extract were analysed by Western blot (section 2.3.3).

2.8.9. Purification of *Strep*-tagged PrimPol from human cultured cells

2.8.9.1 Small scale affinity purification

For small scale affinity purification of soluble *Strep*-tagged PrimPol from human cultured cells, Flp-In T-REx-293 cells engineered for inducible expression of PrimPol-FLAG-*Strep* were seeded in two 175 cm^2 flasks and one day before they became confluent, 10 ng/ml doxycycline was added to one of the flasks inducing PrimPol expression. Cells were harvested by resuspension in PBS and collected by centrifugation (1500 rpm, 5 minutes), and the non-induced and induced pellets each lysed in 1 ml lysis buffer (150 mM NaCl, 30 mM Tris pH 7.4, 0.5 % NP-40, with Roche protease inhibitor cocktail), and place on a tumbler at 4 ºC for at least 20 minutes. The lysate was cleared by centrifugation (10 000 g, 10 minutes, 4 ºC), 100 µl retained as input, and pre-washed 100 µl *Strep*-tactin (Invitrogen) 1:1 slurry added, before incubation for at least 1 hour on a tumbler at 4 ºC (can leave overnight). The material bound to the *Strep*-tactin resin was washed 3 times by brief centrifugation (30 seconds, 3000 g, 4 ºC), resuspension in an excess of wash buffer (same as lysis buffer except 0.1 % NP-40) followed by at least 5 minutes on a tumbler at 4 ºC. Proteins bound specifically to the *Strep*-tactin resin were eluted with 200 µl 2 mM desthiobiotin (in 150 mM NaCl, 30 mM Tris pH 7.4), 3-5 successive elutions were performed. The resulting samples were analysed by Western blot (section 2.3.3).

2.8.9.2. Affinity purification for mass-spectrometry analysis

For the large scale affinity purification of soluble PrimPol for mass-spectrometry analysis, thirty 175 cm^2 flasks of confluent Flp-In T-REx-293 cells engineered for PrimPol expression were used, 1 day before harvesting, PrimPol expression was induced in 15 of these flasks by addition of 10 ng/ml doxycycline. Following harvesting and collection, cell pellets (~1 g each) were lysed in 15 ml lysis buffer and incubated at 4 ºC on a rocker for 20 minutes. Input was retained (500 µl) and 1 ml of *Strep*-tactin resin (packed volume) added to the lysate and placed on a rocker for 2 hours at 4 ºC. Washes were also performed in batch mode and then the *Strep*-tactin resin was transferred to a gravity flow column and washed further. Five successive 500 µl elutions with 2 mM desthiobiotin were performed, and each snap frozen with 10% glycerol. Following Western blot analysis to determine which affinity purifications were successful, the chosen elutions were concentrated using a VivaSpin 10 000 kDa molecular filter before resolving on a Bis-Tris 4-20 % gel and colloidal Coomassie staining (Invitrogen). Whole lane gel

extraction was performed with each lane being divided into 1-2 mm bands which were placed in a 96 well plate before trypsin digestion and mass-spectrometry analysis (in collaboration with Mark Sekel, CRUK, Clare Hall).

2.8.9.3. Treatment of purified PrimPol with lambda phosphatase

PrimPol was affinity purified using *Strep*-tactin resin as detailed above, except for the addition of PhosSTOP phosphatase inhibitor cocktail (Roche) in the lysis buffer. Following 3 washes in wash buffer, the material bound to the *Strep*-tactin resin was washed with protein metallophosphatase (PMP) buffer (New England Biolabs), and then divided between 3 clean eppendorf tubes. To each, PMP buffer supplemented with 1 mM $MnCl_2$ was added with either 400 U Lambda Protein Phosphatase (New England Biolabs) and PhosSTOP, just the Lambda Protein Phosphatase, or neither of the two. Following a 20 minutes incubation at 30 ºC with agitation, the resin was collected by centrifugation and boiled in Laemmli sample buffer before analysis by Western blot (section 2.3.3).

2.9. Yeast two-hybrid methods

The Two-Hybrid Matchmaker 2 System (Clontech) using the *S. cerevisiae* Y190 strain was used, in conjunction with pGBKT7 (containing the GAL4 DNA binding domain) and pACT2 (containing the GAL4 DNA activation domain) vectors. Plasmid constructs used in the yeast two-hybrid assay, and the primers used to generate these constructs, are detailed in Tables 2.10 and 2.11 respectively.

2.9.1. Yeast culture

Y190 was grown in Yeast Extract medium (YE; 0.5 % (w/v) yeast extract, 3 % (w/v) glucose, 0.02 % (w/v) adenine, and 0.1 % (w/v) uracil, histidine, arginine, and leucine) or on YEA plates (YE medium solidified with 2.5 % granulated agar) at 30 ºC. For short term storage plates of yeast were stored at 4 ºC, however for long term storage glycerol stocks were made. A single colony was resuspended in growth medium (YE) and 50 % glycerol added to a final concentration of 25 % before transferring to -80 ºC. To restart a culture, a small proportion of the glycerol stock was streaked onto plates (YEA) and incubated at 30 ºC for 3-5 days.

2.9.2. Yeast transformation

One millilitre of YE was inoculated with a few colonies of Y190, vortexed vigorously to disperse clumps, and transferred to 50 ml of YE to be incubated overnight at 30 ºC with shaking (250 rpm) until stationary phase (OD_{600} > 1.5). The culture was then diluted to an

Table 2.10. Plasmid constructs for yeast two-hybrid

Construct (vector:insert)	Source	Purpose	Construction
pGBKT7	Clontech	Yeast two-hybrid, binding domain	
pGBKT7:PrimPol	Andrew Green, Doherty Lab	Express PrimPol-BD fusion	NdeI - SalI/XhoI
pGBKT7:PrimPol 1-341	This thesis	Express PrimPol 1-341-BD fusion	SDM - Stop codon 341
pGBKT7:PrimPol 370-560	This thesis	Express PrimPol 371-560-BD fusion	Inverse PCR - delete residues 2-369
pPACT2	Clontech	Yeast two-hybrid, activation domain	

Constructs used and generated in this thesis for expression of proteins in *S.* cerevisiase for yeast two-hybrid analysis. 'Construction' column contains either endonuclease restriction sites that the gene of interest is cloned between, or amino acid residues mutated or deleted by site-directed mutagensis (SDM) PCR or inverse PCR respectively. pGBKT7 allows expression of protein-fusions with the GAL4 DNA binding domain (BD), whilst pACT2 allows expression of protein-fusions with the GAL4 transcriptional activation domain. PrimPol refers to human gene coiled-coil domain containing protein (CCDC)111.

Table 2.11. Primers used to generate plasmid constructs for yeast two-hybrid analysis

Name	Sequence 5'-3'	Construct (vector:insert)	Notes
CCDC_341STOP_F	ggttctcagatact**TGA**cgaattcttac	pGBKT7:PrimPol 1-341	
CCDC_341STOP_R	gtaagaattcg**TCA**agtatctgagaacc		
CCDC_371-end_F	GGTTTTCAGTGTTCTCCCTATCC	pGBKT7:PrimPol 370-560	Primers contain 5' phosphate modification for inverse deletion PCR
CCDC_371-end_R_pGBKT7	CATATGCAGGTCCTCCTCTG		

Primers designed and used in this thesis to generate plasmid constructs for yeast two-hybrid analysis. Sequences of mutated bases in primers are in bold. 'Construct' column refers to the construct generated using the primer sets. pGBKT7:PrimPol 1-341 expresses PrimPol residues 1-341 as a fusion to the GAL4 DNA binding domain, whilst pGBKT7:PrimPol 371-560 expressed residues 370-560 as a binding domain fusion. PrimPol refers to human gene coiled-coil domain containing protein (CCDC)111.

OD_{600} of 0.05 in 100 ml of YE and returned to the incubator for a further 3-4 hours until an OD_{600} of 0.2 was reached. The cells were collected by centrifugation (2500 rpm, 5 minutes) and resuspended in 10 ml sterile TE (10 mM Tris pH 7.5, 1 mM EDTA) before centrifugation. The cells were then resuspended in 0.5 ml freshly prepared sterile LiAc-TE (0.1 M lithium acetate, 10 mM Tris pH 7.5, 1 mM EDTA) and kept on ice.

To 5 µg of each plasmid (pGBKT7 and pACT2) and 50 µg carrier DNA (herring sperm DNA denatured at 90 ºC for 20 minutes), 100 µl competent yeast cells were added and mixed well by vortexing. To this mixture, 600 µl of sterile PEG LiAc-TE (40 % (w/v) PEG 3350 in LiAc-TE) was added and vortexed at high speed for 10 seconds. Following a 30 minutes incubation at 30 ºC with shaking, 70 µl 100 % DMSO was added and mixed by inversion, the mixture heat-shocked at 42 ºC for 15 minutes, before being chilled on ice for a further 1-2 minutes. The cells were collected by centrifugation (3000 rpm, 5 minutes), resuspended in 500 µl TE, and 100 µl plated on yeast minimal medium plates (YMM; yeast nitrogen base (YNB), 10 % glucose, 2.5 % granulated agar) supplemented with 20 mg/ml adenine and histidine (to select for pGBKT7 and pACT2). Colonies grew over 2-3 days in a 30 ºC incubator.

2.9.3. Detection of interaction

Colonies of yeast cells transformed with both pGBKT7 and pACT2 were picked and resuspended in 20 µl sterile TE, vortexed to disperse clumps. Half was spread on YMM plates supplemented with adenine and histidine and the other half on YMM plates supplemented with adenine and 25 mM 3-amino-1, 2, 4-triazole (3-AT), a competitive inhibitor of the yeast HIS3 protein therefore inhibiting low-level leaky expression (Durfee *et al*., 1993). At least 4 colonies were streaked per plate. Following 2-3 days at 30 ºC, the growth of cells in the absence of histidine (3-AT plates) indicated a potential interaction of the proteins cloned in pGBKT7 and pACT2. A liquid X-Gal assay was also performed to detect a possible interaction; first, the cells re-streaked on YMM plates containing adenine and histidine were resuspended in 150 µl Z buffer (60 mM Na_2HPO_4, 60 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, pH 7) and lysed by 3 cycles of freeze-thawing in liquid nitrogen. To this, 150 µl Z buffer containing X-Gal and β-mercaptoethanol was added and colour was allowed to develop at 30 ºC. If the mixture turned blue this indicated a potential interaction.

2.10. Bioinformatics

Gene sequences were retrieved via the National Centre for Biotechnology Information (NCBI) (Sayers *et al*, 2009) from GenBank (Benson *et al*, 2009). Typanosomatid gene

sequences were retrieved from TriTryp database (Aslett *et al*, 2009). Multiple sequence alignments were generated using ClustalW2 (Larkin *et al*, 2007), TCoffee (Notredame *et al*, 2000), MUSCLE (Edgar, 2004), and MAFFT (Katoh *et al*, 2009) in the alignment editor JALview (Waterhouse *et al*, 2009). Phylogenetic trees were generated using ClustalW2 and edited using iTOL (interactive tree of life; Letunic and Bork, 2011). ProtParam (Gasteiger *et al*, 2005) was used for computation of physical and chemical parameters of proteins, and various other programs on the ExPASy server (Gasteiger *et al*, 2003).

Chapter 3

The African trypanosome contains two PrimPol-like proteins and one is essential for completion of DNA replication

3.1. Introduction

Eukaryotic genomes encode numerous DNA polymerases required for the replication and maintenance of the genome (Lange *et al*., 2011). In contrast, with the exception of trypanosomatid kinetoplast DNA replication, only a single DNA primase has been characterised to date, being Prim1 of the Pol α-Prim complex. Prim1 is the reference member of the AEP superfamily and is essential for the initiation of DNA replication in all eukaryotes analysed so far (Muzi-Falconi *et al*., 2003). In archaea and prokaryotes, AEPs have been shown to be versatile polymerases capable of a number of nucleotidyl transferase activities, and in some bacteria function in DNA break repair (Lao-Sirieix *et al*., 2005; Brissett and Doherty, 2009). Currently unpublished work in the laboratories of Aidan Doherty and collaborator Luis Blanco (CBMSO, Madrid) has identified an additional eukaryotic AEP-like enzyme, the human protein CCDC111, which was previously reported to be a putative member of the AEP superfamily (Iyer *et al*., 2005). Preliminary experiments have shown that recombinant human CCDC111 is a versatile enzyme *in vitro*, being capable of both primase and polymerase activities, and so the protein was renamed PrimPol. To complement the characterisation of human PrimPol (Chapters 5-7) it was decided to study PrimPol in a relatively 'simpler' eukaryotic model organism. The African trypanosome was chosen (for reasons later explained) and this is the topic of the current chapter.

T. brucei, the African trypanosome, along with the closely related *T. cruzi* and *Leishmania*'s, are protozoan parasites of significant medical importance (World Health Organisation, 2008). *T. brucei* is the causative agent of human African trypanosomiasis, a typically fatal condition that is endemic in sub-Saharan Africa (Brun *et al*., 2010). In addition to their medical importance, trypanosomatids also present a unique evolutionary perspective, as they are among the earliest diverging organisms from the eukaryotic tree (Simpson *et al*., 2006). Replication of the nuclear genome in trypanosomatids is less well understood than replication of its mitochondrial counterpart, although current knowledge is advancing (Tiengwe *et al*., 2012b), and all the major components of the higher eukaryotic replication machinery are present in trypanosomatid genomes suggesting similar mechanisms may exist (El-Sayed *et al*., 2005). Most conventional DNA repair pathways appear to operate in trypanosomatids (Passos-Silva *et al*., 2010), however, consistent with early divergence, the replication origin licensing machinery resembles that of archaea (Godoy *et al*., 2009; Tiengwe *et al*., 2012a) and the nonhomologous end-joining break repair pathway does not appear to operate (Burton *et al*., 2007; Glover *et al*., 2008).

3.2. Trypanosomatid genomes encode two PrimPol-like proteins

To identify PrimPol homologues a Position-Specific Iterated (PSI)-BLAST search (Altchul *et al*., 1997) was performed using the human PrimPol (CCDC111) sequence as query. As previously reported (Iyer *et al*., 2005), PrimPol homologues were readily identified in a broad range of multicellular and unicellular eukaryotes, including animals, plants, and protists. PrimPol was not identified in some animals such as *Drosophila* and *C. elegans*, and was only identified in one fungus, the parasitic *B. dendrobatidis*. Notably, a second group of lower scoring but significant hits were identified in the genomes of trypanasomatids, which were not previously reported by Iyer and colleagues (2005). The sequences of a range of PrimPol homologues were aligned (see Appendix for full alignment) and subjected to phylogenetic analysis, which confirmed that this second group of proteins were divergent members of the PrimPol family (Figure 3.1). As a result, this group of PrimPol homologues were named PrimPol2 and the previously identified trypanosomatid PrimPol homologues were named PrimPol1 (Figure 3.1). As trypanosomatids are amongst the earliest diverging eukaryotic organisms that are readily amenable to genetic manipulation, they appear to represent an excellent model system to characterise these novel enzymes.

3.2.1. PrimPol1 and PrimPol2 in the African trypanosome *T. brucei*

The PrimPol homologues of the African trypanosome, *T. brucei*, were termed TbPrimPol1 (Tb927.5.4070) and TbPrimPol2 (Tb927.10.2520). TbPrimPol1 is 608 amino acids in length and has ~16 % identity with human PrimPol, which is slightly shorter at 560 amino acids. TbPrimPol2 is 732 amino acids long and shares ~11 % identity with its human counterpart. Despite their sequence divergence from human PrimPol and from each other, sharing ~10 % identical residues, both TbPrimPol1 and 2 contain the characteristic domains of the PrimPol family (Figure 3.2). The TbPrimPol1 and 2 amino-termini contain the catalytic AEP domain comprised of three highly conserved motifs: motif I is hhhDhE (where 'h' is a hydrophobic residue), motif II SxH, and motif III hD (Figure 3.3a). Structural and site-directed mutagenesis studies of AEPs has demonstrated that these motifs are essential for catalysis; motif I and III are required for binding of divalent metal ions and motif II is required for nucleotide binding (Copeland and Tan, 1995; Augustin *et al*., 2001; Ito *et al*., 2003; Lao-Sirieix and Bell, 2004). Unlike other PrimPol family members, TbPrimPol2 does not contain the variation in motif I of the catalytic AEP domain (DxE), which is also shared by the archaeal pRN1-PrimPol (Lipps *et al*., 2003), but rather resembles the majority of AEP family members, containing DxD (Figure 3.3a). Carboxyl-terminal to the AEP domain, TbPrimPol1 and 2 also contain the CHC_2 zinc finger motif with homology to the herpesviral UL52 primase (Figure 3.3b).

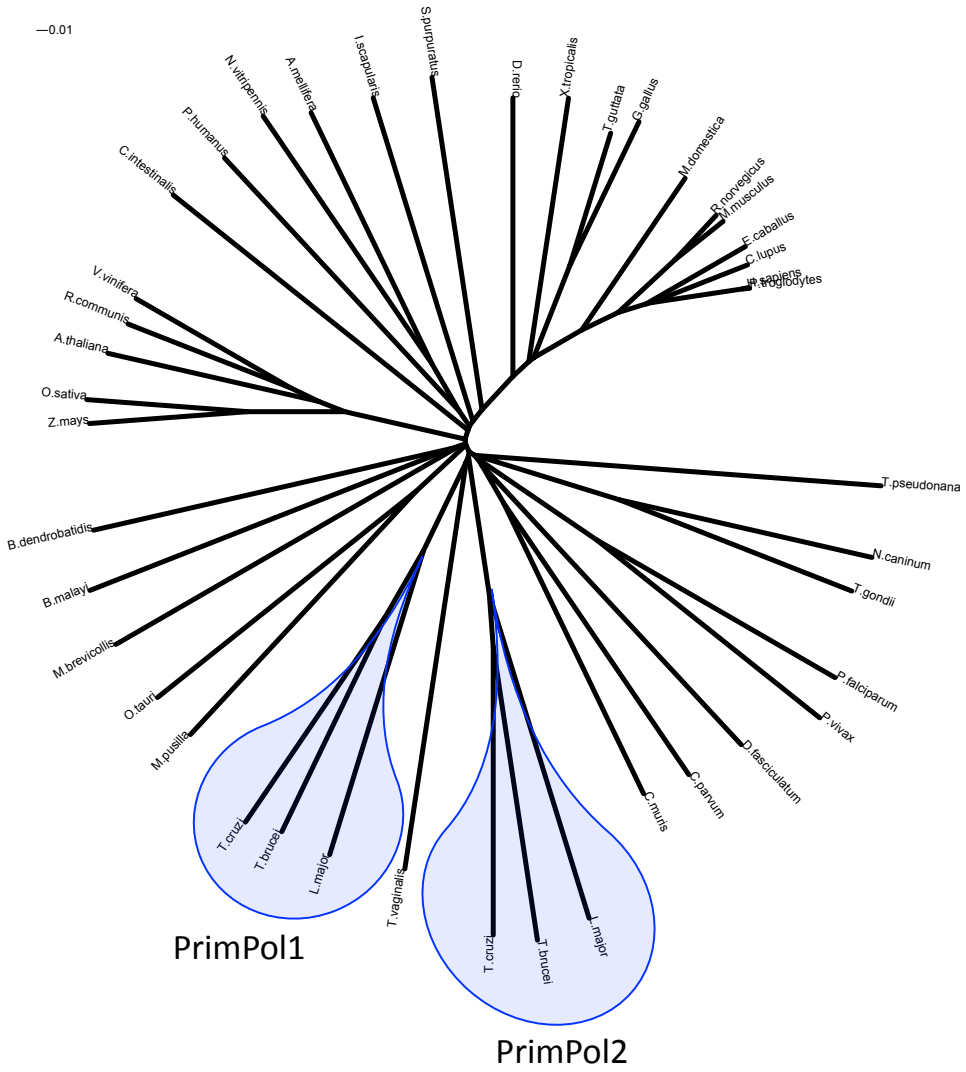


Figure 3.1. Phylogenetic analysis of PrimPol family members.
PrimPol homologues were identified by PSI-BLAST searching (Altchul *et al*., 1997) the available eukaryotic genomes using human PrimPol and lower eukaryotic PrimPol homologues as queries. A multiple sequence alignment was assembled with 42 of the homologues identified, representing a broad range of homologues from animals, plants, and protists (see Appendix for alignment and accession numbers of homologues). Alignments were generated using ClustalW2 (Larkin *et al.*, 2007), TCoffee (Notredame *et al.*, 2000), MUSCLE (Edgar, 2004) and edited in JALview (Waterhouse *et al*, 2009). The phylogenetic tree was generated using ClustalW2 and annotated using iTOL (Letunic and Bork, 2011). The two trypanosomatid PrimPol families, PrimPol1 and PrimPol2, are indicated.

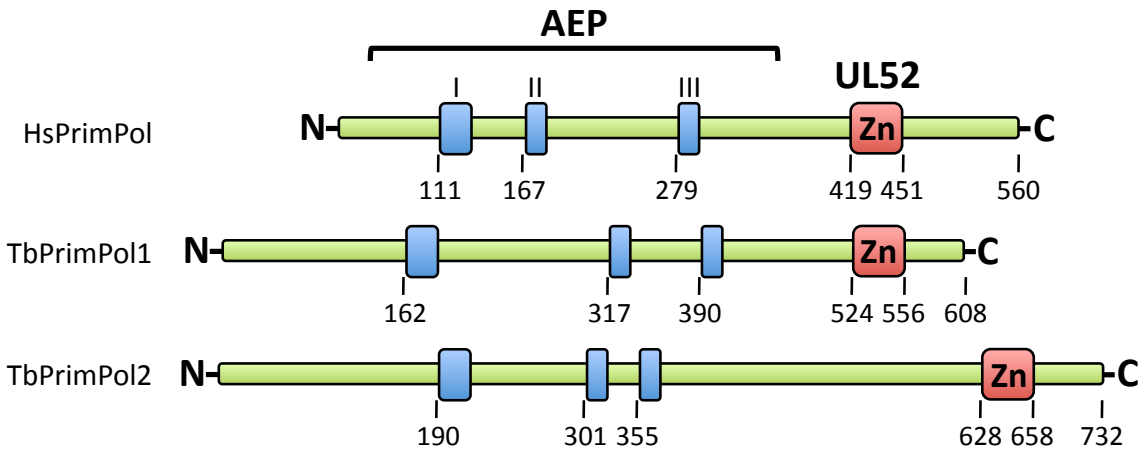


Figure 3.2. Conserved motifs and domains of human PrimPol and *T. brucei* homologues. Schematic showing the locations of conserved domains and motifs in human PrimPol and TbPrimPol1 and 2. The conserved domains were identified by generating a multiple sequence alignment containing PrimPol homologues from a broad range of multicellular and unicellular eukaryotes (see Figure 3.1, and Appendix for full alignment). The locations of the catalytic AEP motifs I, II, and III, and the herpesviral UL52 zinc finger are indicated.

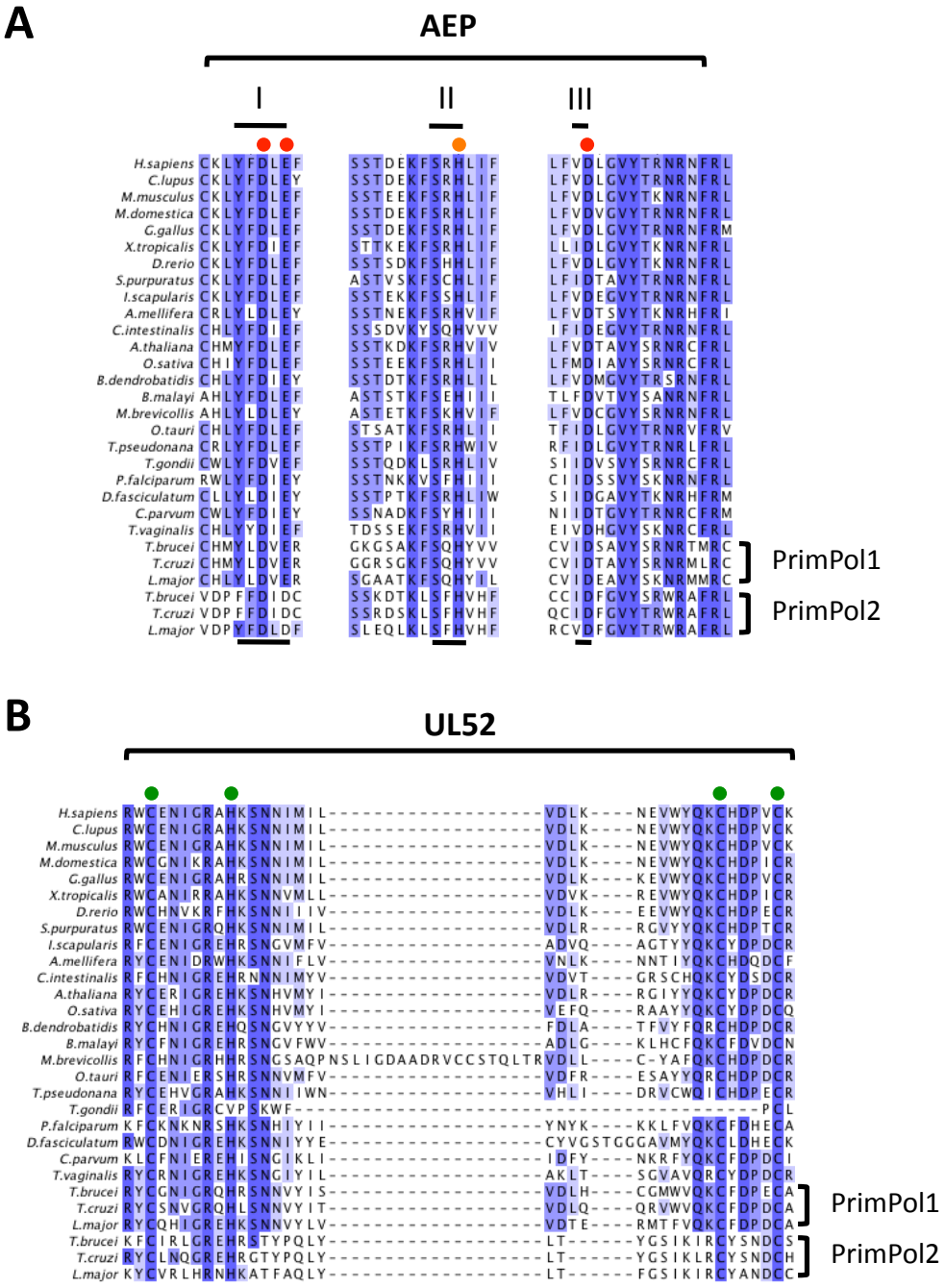


Figure 3.3. Sequence alignments of the conserved AEP and UL52 motifs from a broad range of PrimPol homologues. A multiple sequence alignment was generated using PrimPol homologues from a broad range of multicellular and unicellular eukaryotes (see Figure 3.1, and Appendix for full alignment and accession numbers). Motifs of the catalytic AEP domain (**A**) are shown with the essential residues indicated. Residues in motifs I and III (red circles) are predicted to be required for divalent metal ion binding and motif II (orange circle) for nucleotide binding. The herpesviral UL52 CHC_2 zinc finger (**B**) is also shown. Residues predicted to be essential for chelation of zinc are indicated (green circles). The trypanosomatid PrimPol1 and PrimPol2 families are also indicated. Figure produced using JALview (Waterhouse *et al*, 2009). Blue shading indicates ≥ 40 % sequence conservation.

Although TbPrimPol1 and 2 are divergent from higher eukaryotic PrimPol's they are relatively well conserved among trypanosomatids; the *T. cruzi* and *L. major* homologues share 50 and 36 % identity with their *T. brucei* PrimPol1 counterpart, and 44 and 34 % with their PrimPol2 counterpart respectively (Figures 3.4 and 3.5).

3.3. Generation of strains to study the cellular roles of TbPrimPol1 and 2

In order to explore the cellular roles of TbPrimPol1 and 2 in the African trypanosome, tetracycline-inducible RNAi strains and *in situ* epitope-tagged strains were first generated for each of these novel enzymes. These were made in bloodstream form trypanosomes (the pathogenic life cycle stage) specifically the Lister 427 strain (MiTat 1.2, clone 221a) and its transgenic derivative the 2T1 strain (Alsford *et al*., 2005), using constructs developed by Alsford and Horn (2008).

3.3.1. TbPrimPol1 and 2 inducible RNAi strains

To generate inducible RNAi strains (overview in Figure 3.6), first a DNA target sequence of 400-600 bases was selected from each of the TbPrimPol1 and 2 genes using RNAit software (Redmond *et al*., 2003). Selected sequences will have little sequence similarity to other genes within the *T. brucei* genome and therefore should specifically target either TbPrimPol1 or 2 mRNA. A 401 base pair target sequence was selected for TbPrimPol1 that corresponds to bases 1292-1693 of the 1827 base pair open reading frame. For TbPrimPol2 a 532 base pair target sequence was selected, corresponding to bases 859-1391 of the 2199 base pair open reading frame. These target sequences were amplified from genomic DNA (see Table 2.4 for primers) and cloned individually in a head-to-head conformation in the inducible RNAi vector pRPa^{iSL}, generating constructs pRPa^{iSL}:TbPrimPol1 and pRPa^{iSL}:TbPrimPol2 (Table 2.3). These constructs were linearised and transfected into the 2T1 strain, which in addition to constitutively expressing the Tet repressor, also contains a tagged ribosomal RNA locus for targeted integration (Alsford *et al*., 2005). This targeted integration eliminates variable expression and position effects that can be observed following random integration within the genome (Alsford *et al*., 2005). Correctly integrated transformants were selected by antibiotics and should have tetracycline-inducible expression of stem-loop RNA to target the destruction of either TbPrimPol1 or 2 mRNA, thus knocking down protein levels. Two inducible RNAi clones were generated for TbPrimPol1, and 3 inducible RNAi clones were generated for TbPrimPol2.

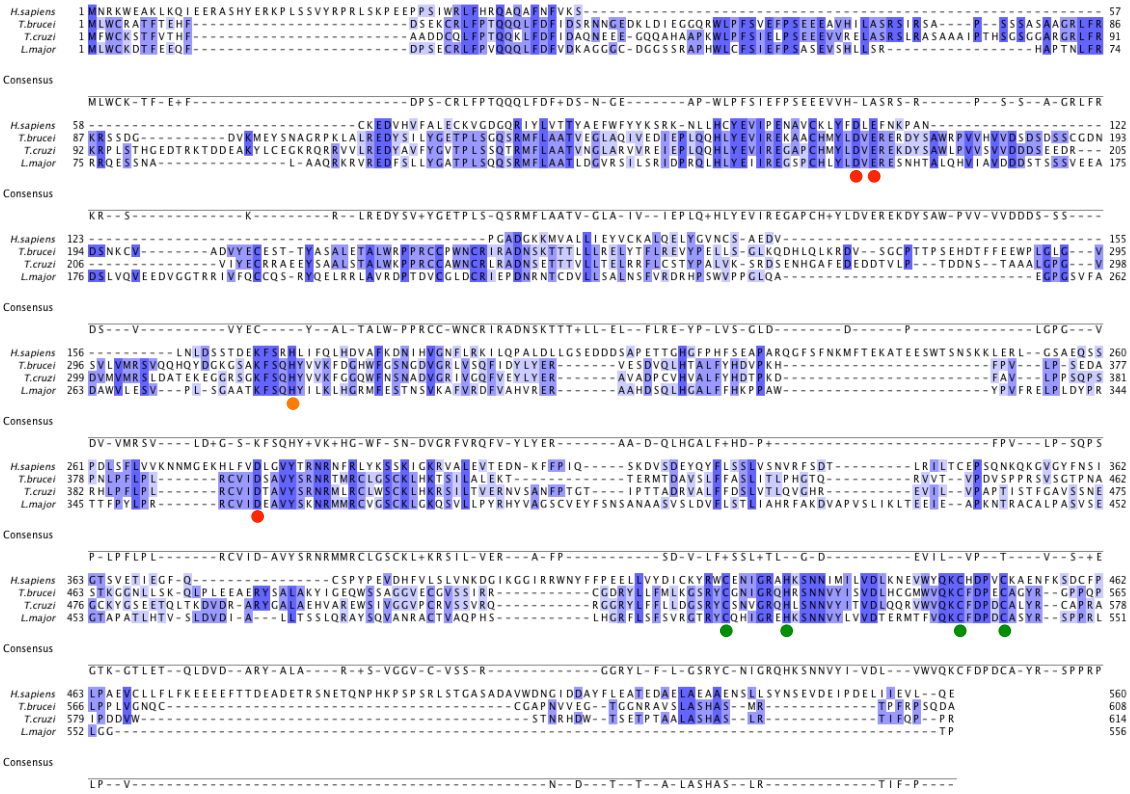


Figure 3.4. Alignment of the trypanosomatid PrimPol1 family.
Multiple sequence alignment of human PrimPol (NP_689896) and the trypanosomatid PrimPol1 family, including *T. brucei* (XP_845077), *T. cruzi* (XP_812815), and *L. major* (XP_001685728). Predicted essential residues in the AEP domain are indicated, residues in motifs I and III (red circles) are required for divalent metal ion binding and motif II (orange circle) for nucleotide binding. Predicted residues essential for chelation of zinc in the UL52 CHC_2 zinc finger are also indicated (green circles). Alignment was generated from a multiple sequence alignment of PrimPol homologues from a broad range of multicellular and unicellular eukaryotes (see Figure 3.1, and Appendix for full alignment). Figure produced using JALview (Waterhouse *et al*, 2009). Blue shading indicates ≥ 40 % sequence conservation.

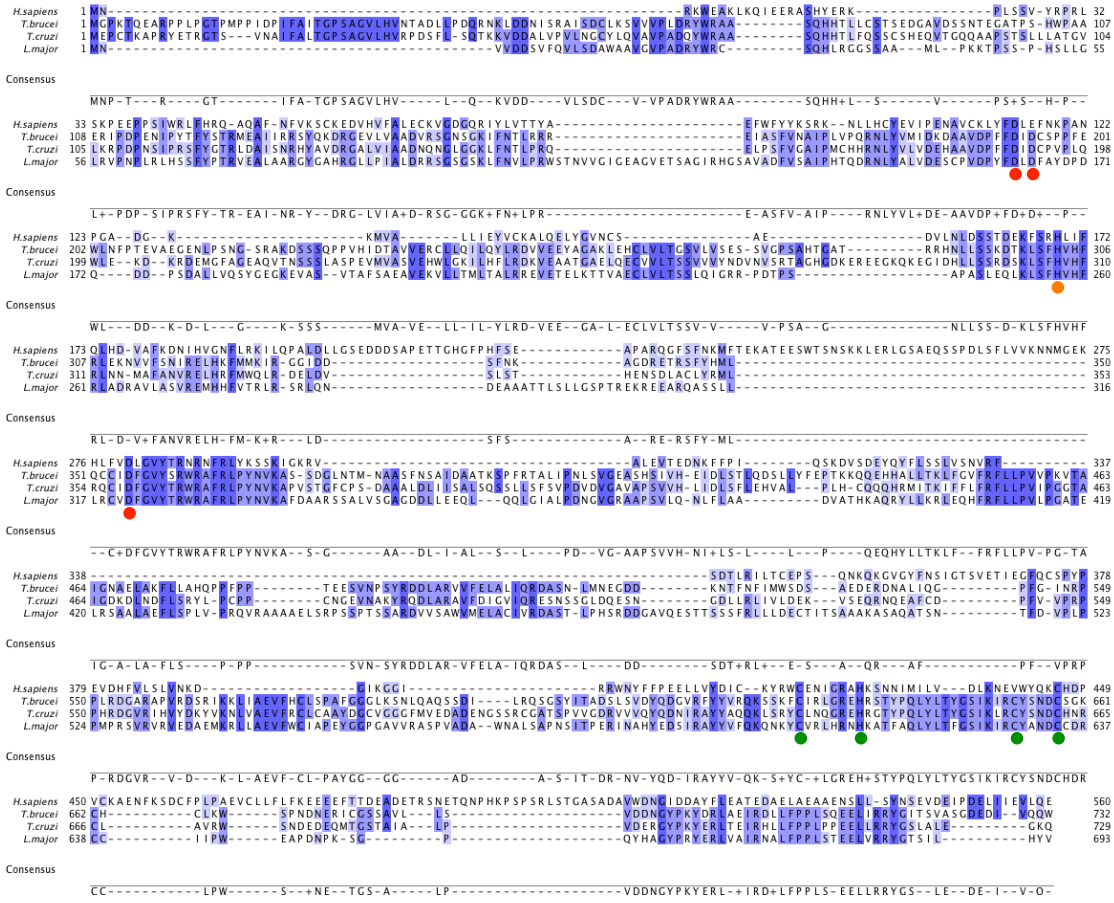


Figure 3.5. Alignment of the trypanosomatid PrimPol2 family.
Multiple sequence alignment of human PrimPol (NP_689896) and the trypanosomatid PrimPol2 family, including *T. brucei* (XP_822505), *T. cruzi* (XP_819101), and *L. major* (XP_001686100). Predicted essential residues in the AEP domain are indicated, residues in motifs I and III (red circles) are required for divalent metal ion binding and motif II (orange circle) for nucleotide binding. Predicted residues essential for chelation of zinc in the UL52 CHC_2 zinc finger are also indicated (green circles). Alignment was generated from a multiple sequence alignment of PrimPol homologues from a broad range of multicellular and unicellular eukaryotes (see Figure 3.1, and Appendix for full alignment). Figure produced using JALview (Waterhouse *et al*, 2009). Blue shading indicates ≥ 40 % sequence conservation.

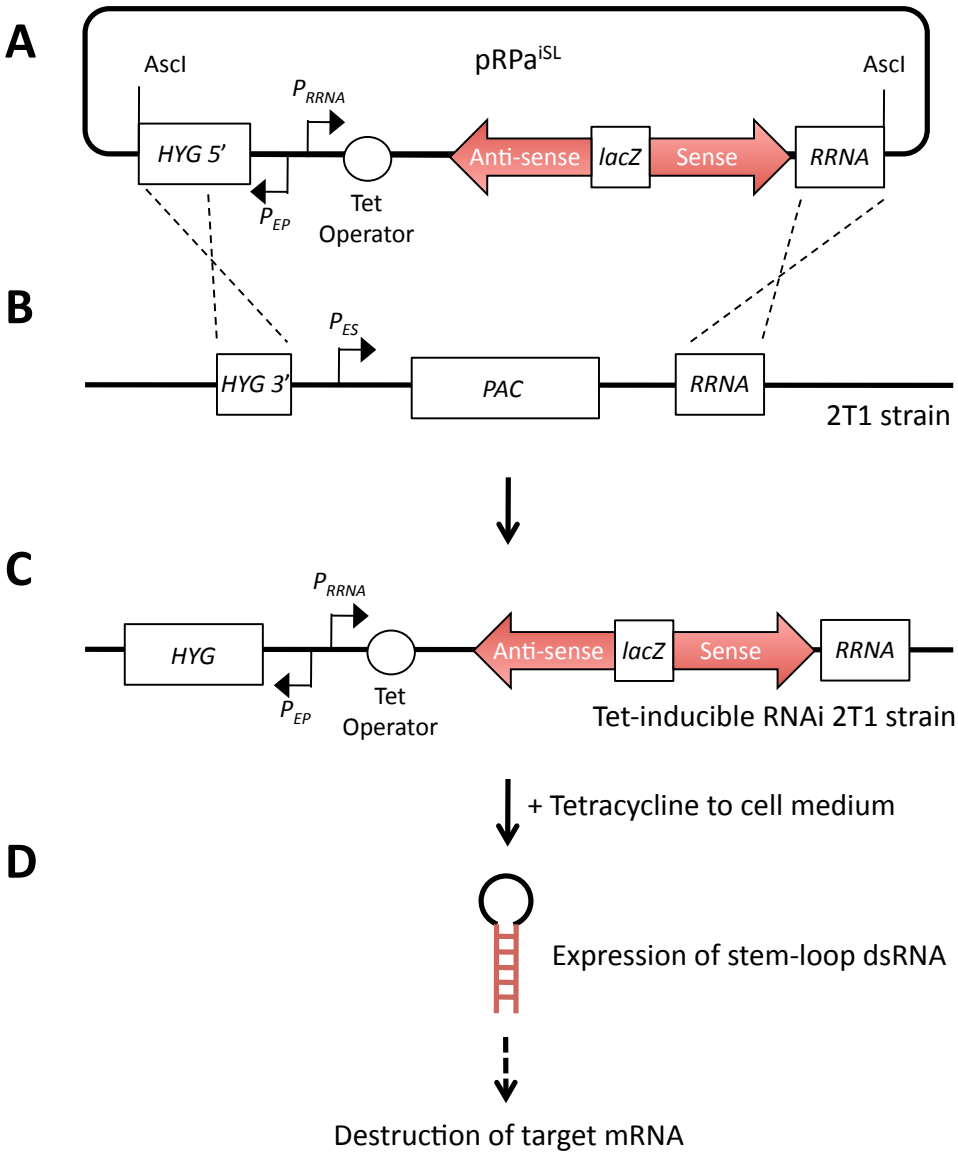


Figure 3.6. Generation of a tetracycline-inducible RNAi strain in *T. brucei*.
(**A**) A 400-600 bp DNA fragment corresponding to a unique sequence from the gene of interest is amplified from genomic DNA and cloned sequentially in a head-to-head conformation (red shading) in the vector pRPa^{iSL}. (**B**) Digestion with AscI gives a construct with terminal *HYG* 5' and *RRNA* spacer sequences for targeting the *HYG* 3'::*PAC*::*RRNA* locus on chromosome 2a in 2T1 cells (*PAC*, puromycin resistance). These cells also express the *Tn10* Tet repressor (*TetR*::*BLE*::*TUB*) (*BLE*, bleomycin resistance). (**C**) Correct integration produces an intact *HYG* ORF and deletes the *PAC* ORF (*HYG*, hygromycin resistance). Transcription of *HYG* is constitutive (P_{EP}, procyclin promoter) allowing selection of stable transformants, and expression of the stem loop RNA is under the tetracycline regulated promoter (P_{RRNA}, RRNA promoter). (**D**) Addition of tetracycline to the culture media induces expression of the stem loop dsRNA which is the trigger for RNAi, knocking down the protein of interest (Alsford *et al*., 2005; Alsford and Horn, 2008).

3.3.2. TbPrimPol1 and 2 *in situ* tagged strains

For the purpose of determining the sub-cellular localisation of TbPrimPol1 and 2, and for monitoring protein level following RNAi induction, *in situ* epitope tagged strains were generated (overview in Figure 3.7). This is the addition of an epitope tag to one of the chromosomal copies of TbPrimPol1 or 2, which is facilitated by the high frequency of homologous recombination of tagging cassettes into target genes observed in *T. brucei* (Lee *et al*., 1990; ten Asbroek *et al*., 1990; Eid and Sollnerwebb, 1991). DNA fragments of 977 and 645 base pairs corresponding to the 3' terminus of the TbPrimPol1 and 2 genes respectively, were amplified from genomic DNA (see Table 2.4 for primers) without their stop codon and cloned in frame with the coding sequence for 12 consecutive c-Myc epitopes in the tagging vector pNAT^{x12M}, generating constructs pNAT^{x12M}:TbPrimPol1 and pNAT^{x12M}:TbPrimPol2 (Table 2.3). These constructs were linearised and transfected into both wild-type and the cognate inducible RNAi strains. Correct integration will result in the addition of 12 c-Myc epitopes to the carboxyl-terminus of one of the chromosomal copies of TbPrimPol1 or 2, and therefore expression of TbPrimPol1^{Myc} or TbPrimPol2^{Myc} will be under their endogenous transcriptional control. Four *in situ* Myc-tagged clones were generated for TbPrimPol1, including 2 in the cognate inducible RNAi strain, and 3 tagged clones were generated for TbPrimPol2, with one in the cognate RNAi strain.

3.3.3. Confirmation of TbPrimPol1 and 2 *in situ* tagged and RNAi strains

To test expression of recombinant TbPrimPol1 and 2 and RNAi induction, the *in situ* tagged RNAi strains were grown in the presence or absence of tetracycline and the lysates prepared from these cells analysed by Western blot. Natively tagged TbPrimPol1^{Myc} has a predicted molecular weight of 85.6 kDa (68 kDa without tag) but expressed as a species of 100 kDa when visualised by Western blot analysis with an anti c-Myc antibody (Figure 3.8a). This larger species was observed in all tagged clones (not shown) and 24 hours after addition of tetracycline to the culture media of the RNAi cells, a substantial reduction in this species was observed (Figure 3.8a). Natively tagged TbPrimPol2^{Myc} has a predicted molecular weight of 99.4 kDa (82 kDa without tag) but expressed as a larger species of approximately 125 kDa when visualised by Western blot analysis (Figure 3.8b). This species was detected in all tagged clones (not shown) and just 8 hours after addition of tetracycline to the culture media of the RNAi cells, the specific depletion of this species was observed (Figure 3.8a). This confirms the successful *in situ* tagging and RNAi-mediated knockdown of TbPrimPol1 and 2.

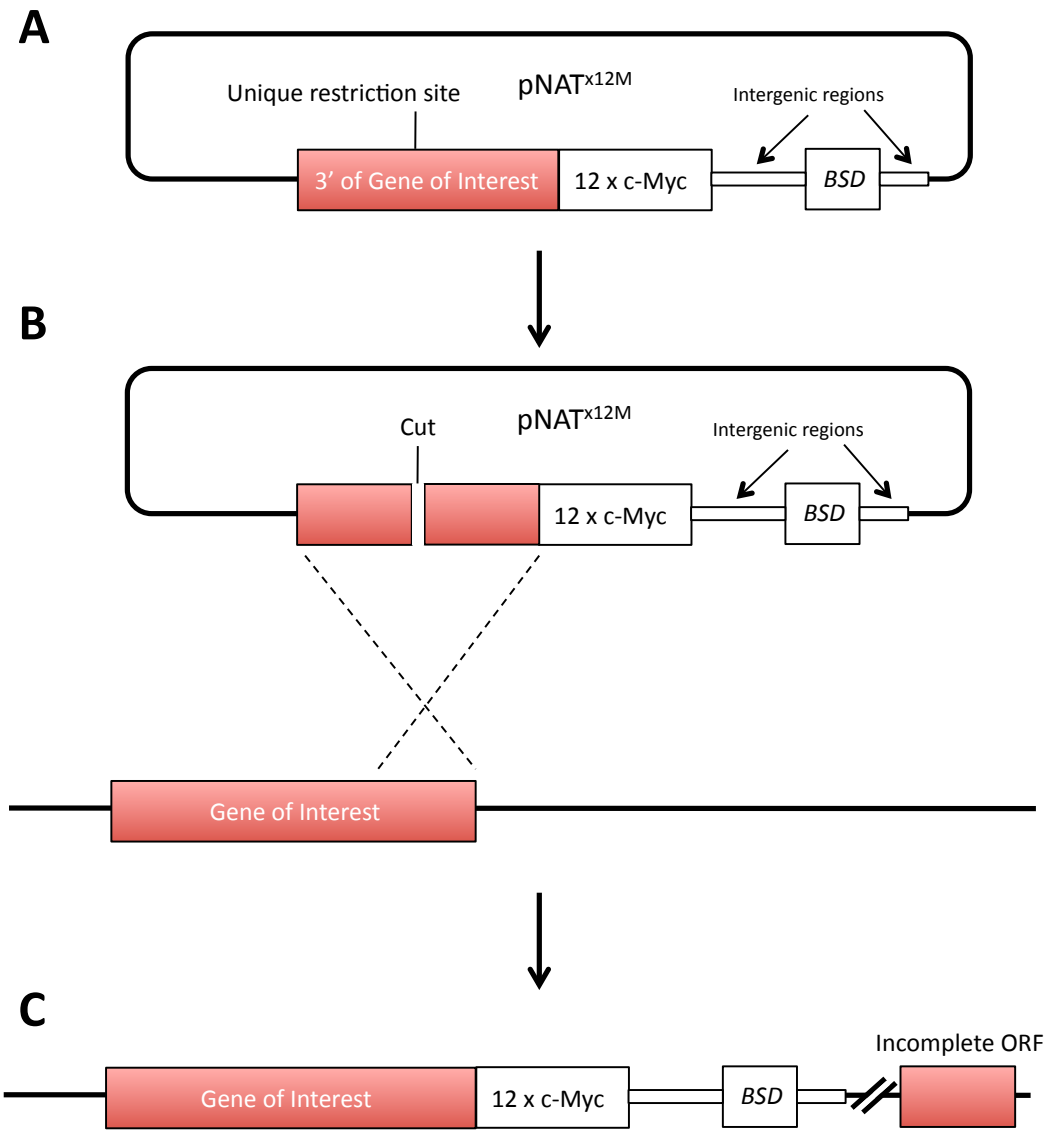


Figure 3.7. Generation of an *in situ* native allele tagging strain in *T. brucei*.
(**A**) A ~1 kb 3' region of the gene of interest is amplified from genomic DNA and cloned into pNAT^{x12M} in frame with 12 consecutive c-Myc epitopes. (**B**) The plasmid is linearised using a unique restriction site within the cloned 3' region creating a construct with terminal regions corresponding to the gene of interest. (**C**) Correct integration will result in the addition of the coding sequence for 12 consecutive c-Myc epitopes to the C-terminus and the 3' UTR will be replaced by *BSD* gene (*BSD*, blasticidine resistance), allowing selection of transformants. The gene of interest now containing a C-terminal c-Myc tag will be under the regulation of its endogenous promoter (Alsford and Horn, 2008).

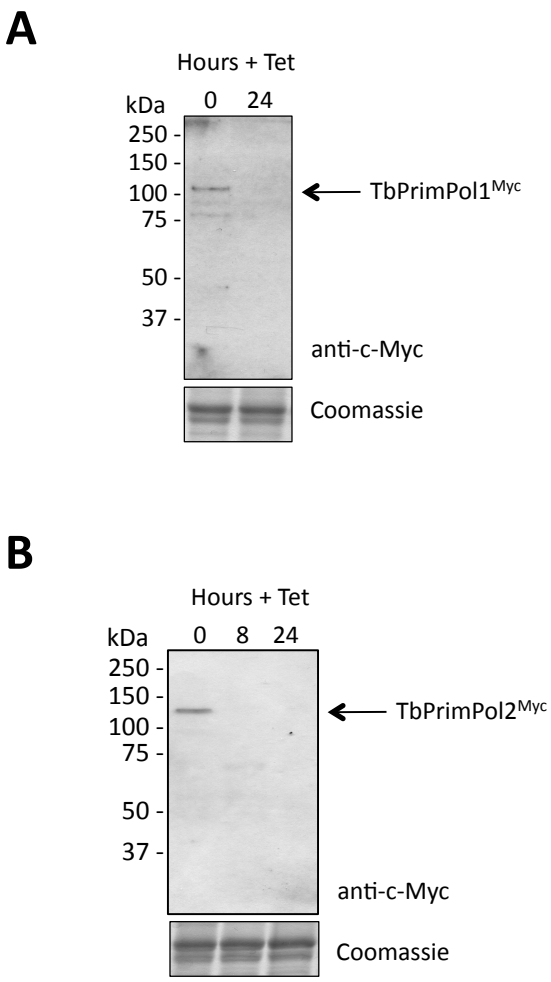


Figure 3.8. Validation of TbPrimPol1 and 2 *in situ* tagged and RNAi strains.
Representative *in situ* tagged TbPrimPol1^{Myc} (**A**) and TbPrimPol2^{Myc} (**B**) inducible RNAi strains were grown in the presence of 1 µg/ml tetracycline for the times indicated and cell lysates prepared and analysed by Western blot with an anti-c-Myc antibody. Lysates were also resolved by SDS-PAGE and Coomassie stained as a loading control.

3.4. TbPrimPol2 is an essential protein in bloodstream form *T. brucei*

It was first investigated whether TbPrimPol1 or 2 were required for cell proliferation. Cultures of RNAi strains were grown in the presence or absence of tetracycline and the cell density monitored over time using a haemocytometer, with the cultures being diluted to their starting density every 24 hours. No proliferative defect was observed following 72 hours TbPrimPol1 RNAi (Figure 3.9a), however, in stark contrast, a severe proliferative defect was observed following TbPrimPol2 RNAi (Figure 3.9b). This defect was clearly visible after 24 hours and was ultimately lethal after 48 hours (Figure 3.9b). Thus, whilst TbPrimPol1 was dispensable for normal cell proliferation in bloodstream form *T. brucei*, TbPrimPol2 was essential. As a result this chapter largely focuses on the essential protein, TbPrimPol2.

3.5. TbPrimPol2 is a nuclear protein

The sub-cellular localisation of TbPrimPol2 was next investigated. The TbPrimPol2 *in situ* Myc-tagged cells were fixed and subjected to immunofluoresence analysis with an anti-c-Myc antibody and DAPI counterstaining, which in the African trypanosome allows visualisation of both the nuclear and mitochondrial kinetoplast genome. TbPrimPol2^{Myc} immunofluoresence was observed in the nuclei of cells (Figure 3.10) suggesting this is where its essential role takes place; however, it was not detected in all cells. The sub-cellular localisation of the dispensable TbPrimPol1 was also investigated but numerous attempts failed to detect any TbPrimPol1^{Myc} immunofluoresence.

3.6. TbPrimPol2 is expressed in G2/M phase cells

Given that TbPrimPol2^{Myc} was only detected in a sub-population of cells, this prompted the question of whether TbPrimPol2 expression was cell cycle regulated. Cell cycle stage can be determined relatively easily in an unperturbed African trypanosome population since the mitochondrial kinetoplast genome is visible by DAPI staining and divides in a cell cycle dependant manner, preceding nuclear mitosis (Figure 3.11a). Thus, a cell with 1 nucleus and 1 kinetoplast (1n1k) represent G1/S, 1 nucleus and 2 kinetoplasts (1n2k) represent G2/M, and 2 nuclei and 2 kinetoplasts (2n2k) represent a post-mitosis but pre-cytokenesis cell (Woodward and Gull, 1990; Siegel *et al*., 2008). TbPrimPol2^{Myc} was detectable by immunofluorescence in ~10 % of cells (Figure 3.11b), with the majority (~7.5 %) being 1n2k cells. Around 1.5 % of cells with detectable TbPrimPol2^{Myc} were 1n1k, and ~1 % were 2n2k. This is a striking result for a DNA synthetic enzyme given that 1n2k cells (which constitute ~15 % of an asynchronous population) are indicative of G2/M phase, and DNA replication is advanced, if not complete, in these cells. This

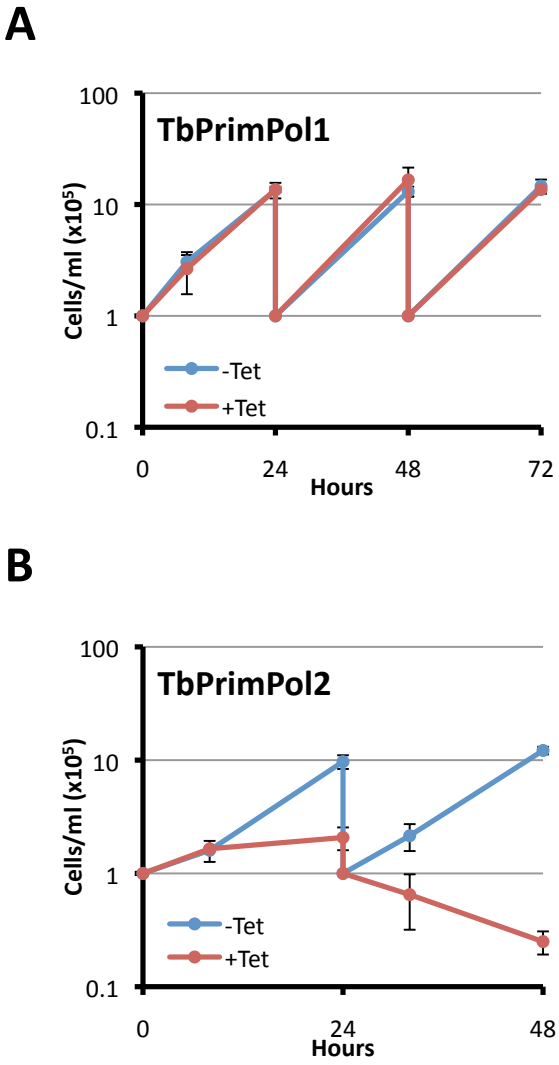


Figure 3.9. TbPrimPol2 is essential whilst TbPrimPol1 is dispensable in bloodstream form *T. brucei*. Representative growth curves of TbPrimPol1 (**A**) and TbPrimPol2 (**B**) inducible RNAi strains grown in the presence (red line) or absence (blue line) of 1 µg/ml tetracycline. The cell density of the cultures was determined using a haemocytometer at 8, 24, 36, 48, and 72 hours, and cultures were diluted to 10^5 cells/ml every 24 hours as appropriate. For TbPrimPol1 RNAi, 3 independent experiments were performed in duplicate, each with independent clones (1 with native allele tagged). For TbPrimPol2 RNAi, 4 experiments were performed in duplicate, each with independent clones (1 with native allele tagged). Error bars represent one standard deviation.

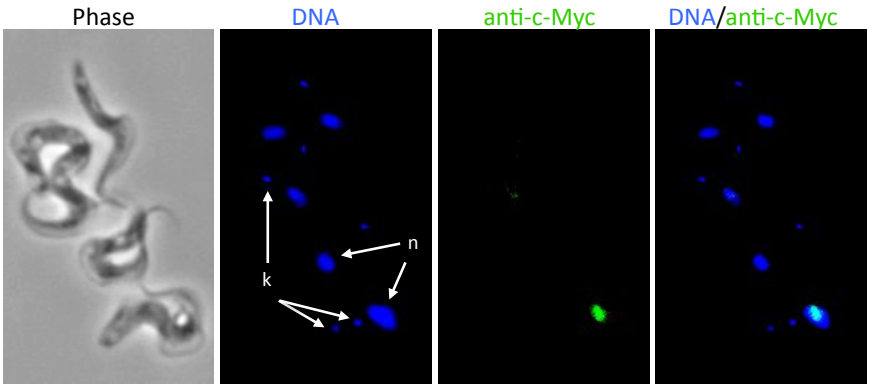


Figure 3.10. TbPrimPol2 is a nuclear protein.
A representative *in situ* tagged TbPrimPol2^{Myc} strain was subjected to immunofluoresence analysis with an anti-c-Myc antibody (green). DNA was counterstained with DAPI, nuclei (n) and the mitochondrial kinetoplast (k) genome are indicated. Images were taken with a 100x objective.

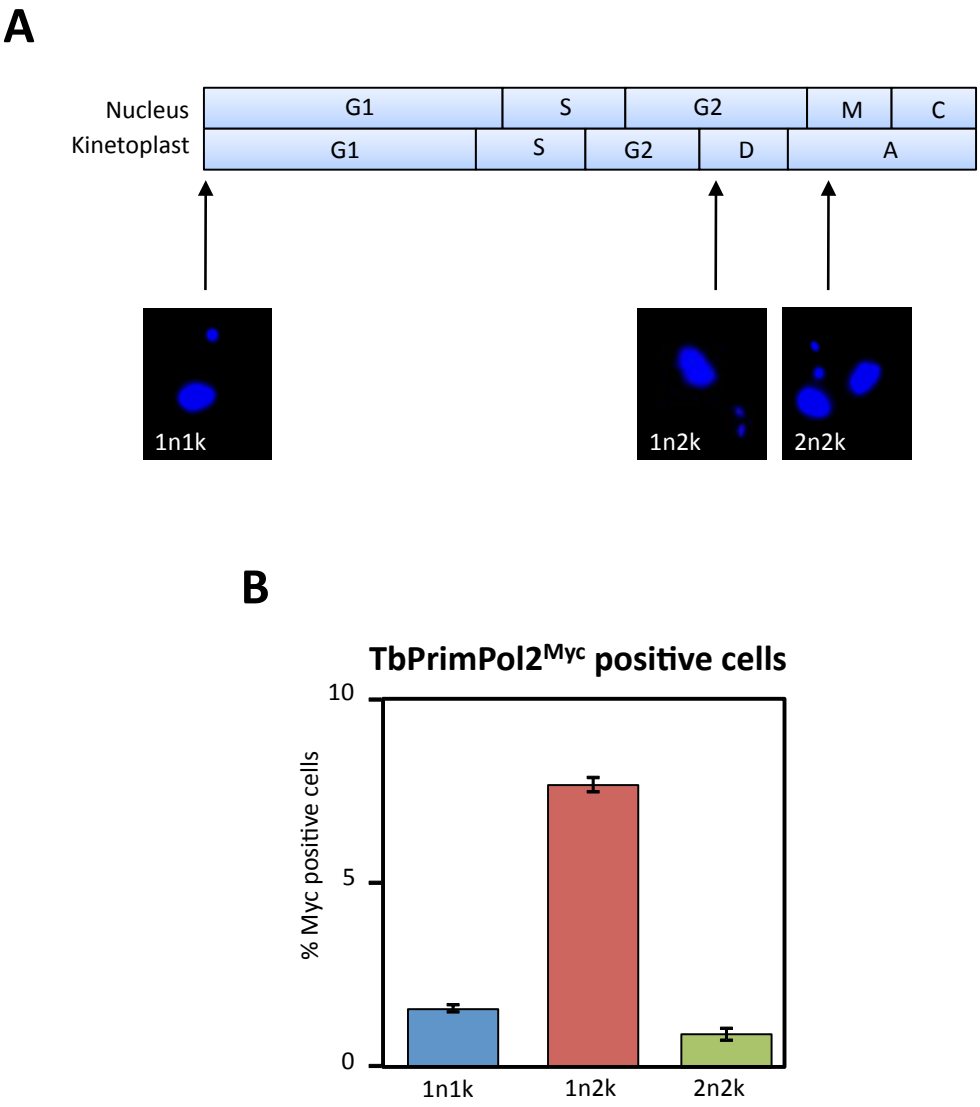


Figure 3.11. TbPrimPol2 is predominantly detected in 1n2k (G2/M) cells.
(**A**) Schematic showing the relative timing of replication and division of the nuclear and mitochondrial (kinetoplast) genome in the *T. brucei* cell cycle. During nuclear G1 and S phase a cell contains 1 nucleus and 1 kinetoplast (1n1k). As kinetoplast division (D) occurs before nuclear mitosis (M) a nuclear G2/M phase cell will have 1 nucleus and 2 kinetoplasts (1n2k). A post-mitotic cell that is yet to undergo cytokinesis (C) will contain 2 nuclei and 2 kinetoplasts (2n2k). G1 and G2 denote growth phases, S is the DNA synthesis phase, and A is apportioning, in which the basal bodies move apart. Figure adapted from Woodward and Gull, 1990. Representative fluoresence DAPI images of each configuration of nuclear to kinetoplast genomes are shown. (**B**) Cells expressing *in situ* tagged TbPrimPol2^{Myc} were analysed by immunofluoresence microcopy using an anti-c-Myc antibody and DAPI counterstaining of DNA. At least 200 cells with TbPrimPol2^{Myc} immunofluorescence were scored for their cell cycle position in 2 independent clones, error bars denote standard deviation.

suggests that the role of TbPrimPol2 may be distinct from that of the canonical DNA primases, required for DNA replication initiation and progression through S-phase. Indeed, the accumulation of TbPrimPol2 at the G2/M phase of the cell cycle suggests a late- or post-DNA replication role.

3.7. TbPri, the putative replicative DNA primase in *T. brucei*

Since TbPrimPol2 is a novel AEP with an essential nuclear function, it was decided to compare TbPrimPol2 and the canonical replicative DNA primase, required for *de novo* synthesis of RNA primers needed to initiate DNA replication. The trypanosomatid genomes encode a putative replicative DNA Primase (El-Sayed *et al*., 2005), which in higher eukaryotes is known to exist as a heterodimer composed of a small catalytic subunit (Prim1) and large accessory subunit (Prim2), which exists associated with Pol α (Muzi-Falconi *et al*., 2003). It was decided to study the large subunit (Tb927.10.3110) that will be referred to as TbPriL (*T. brucei* Primase Large subunit) in keeping with the archaeal nomenclature, as a recently identified kinetoplastid-specific primase was named Pri2 (Hines and Ray, 2011).

3.7.1. Generation of TbPriL *in situ* tagging and inducible RNAi strains

To generate an inducible RNAi strain for TbPriL (overview in Figure 3.6), a 775 base pair target sequence (corresponding to bases 459-1234) was selected from the 1734 base pair open reading frame of TbPriL using RNAit software (Redmond *et al*., 2003). This target sequence was amplified from genomic DNA (see Table 2.4 for primers) and cloned in a head-to-head conformation in the RNAi vector pRPa^{iSL}, generating the construct pRPa^{iSL}:TbPriL (Table 2.3). To generate a TbPriL *in situ* epitope tagged strain (overview in Figure 3.7), a DNA fragment of 754 base pairs corresponding to the 3' terminus of the TbPriL open reading frame was amplified from genomic DNA (see Table 2.4 for primers) without its stop codon, and cloned in frame with the coding sequence for 12 consecutive c-Myc epitopes in the tagging vector pNAT^{x12M}, generating the construct pNAT^{x12M}:TbPriL (Table 2.3). These constructs were used as described previously (section 3.3.1 and 3.3.2) to generate 3 TbPriL inducible RNAi clones and 4 *in situ* Myc-tagged clones, including 2 RNAi clones with a native allele tagged.

3.7.2 Confirmation of TbPriL *in situ* tagged and RNAi strains

To test expression of *in situ* Myc-tagged TbPriL and RNAi induction, the inducible RNAi cells with a native allele tagged were grown in the presence and absence of tetracycline and the lysates prepared from these cells analysed by Western blot. TbPriL^{Myc} has a

predicted molecular weight of 82.8 kDa (65 kDa without tag) but expressed predominantly as a single species of approximately 100 kDa (Figure 3.12). This larger species was observed in all tagging clones (not shown) and 48 hours after the addition of tetracycline to the cell media of the RNAi cells, a substantial reduction of this species was observed (Figure 3.12a). This confirms the successful *in situ* Myc-tagging of TbPriL and RNAi-mediated knockdown.

3.7.3. TbPriL is an essential, nuclear protein in bloodstream form *T. brucei*

It was first investigated whether TbPriL was required for cell proliferation. Cultures of RNAi cells were grown in the presence and absence of tetracycline and the cell density monitored over time as described previously. RNAi knockdown of TbPriL resulted in a severe growth defect, much like TbPrimPol2, that was ultimately lethal after 48 hours (Figure 3.13). Is to be noted that although the growth defect observed in TbPrimPol2 knockdown cells appears more severe than that of TbPriL (Figure 3.9b and 3.13, compare 24 hours), TbPrimPol2 knockdown was much more efficient (Figure 3.8b and 3.12, compare 8 hours) and at a comparable time point, such as 36 hours after RNAi induction, the growth defects look similar. The sub-cellular localisation of TbPriL was next determined using immunofluorescent microscopy. TbPriL^{Myc} was detected in the nuclei of cells (Figure 3.14) as to be expected as a component of the replicative DNA primase. Thus, as expected, TbPrL is an essential, nuclear protein in bloodstream form *T. brucei*.

3.8. TbPrimPol2's cellular role is distinct from canonical DNA primases

It was next investigated whether the severe growth defect that followed TbPrimPol2 and TbPriL knockdown was associated with aberrant cell cycle progression.

3.8.1. Depletion of TbPrimPol2 arrests cells with 1 nucleus and 2 kinetoplasts

As previously described, a useful cytological tool to determine cell cycle phase in *T. brucei* is the division of the mitochondrial kinetoplast genome (Figure 3.11a). Cultures of TbPrimPol2 and TbPriL RNAi strains were grown in the presence of tetracycline and analysed by DAPI staining and fluorescence microscopy to determine the cell cycle distribution. In uninduced cultures, the cell cycle distribution was as expected in an asynchronous population; ~80 % of cells were 1n1k (G1/S), ~15 % were 1n2k (G2/M), and the remainder were 2n2k (post-M). However, following 24 hours knockdown of either TbPriL or TbPrimPol2, the number of 1n1k cells was reduced to 10-20 %, and the number of 1n2k cells increased to 80-90 % (Figure 3.15a and b). The beginning of this shift in cell cycle distribution was visible just 8 hours after TbPrimPol2 RNAi induction

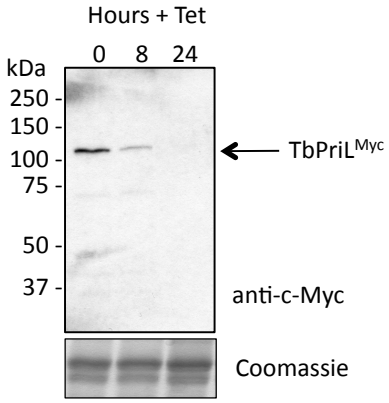


Figure 3.12. Validation of TbPriL *in situ* tagged and RNAi strain.
A representative *in situ* tagged TbPriL^{Myc} inducible RNAi strain was grown in the presence of 1 µg/ml tetracycline for the times indicated and cell lysates prepared and analysed by Western blot with an anti-c-Myc antibody. Lysates were also resolved by SDS-PAGE and Coomassie stained as a loading control.

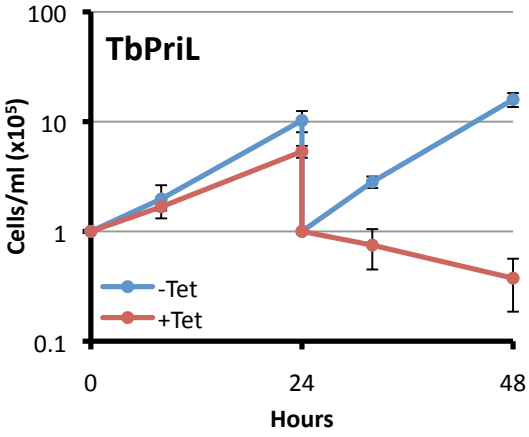


Figure 3.13. TbPriL is essential in bloodstream form *T. brucei*.
Representative growth curve of a TbPriL inducible RNAi strain grown in the presence (red line) or absence (blue line) of 1 µg/ml tetracycline. The cell density of the cultures was determined using a haemocytometer at 8, 24, 36, and 48, and cultures were diluted to 10^5 cells/ml every 24 hours as appropriate. Three experiments were performed in duplicate, each with independent clones (1 with native allele tagged), error bars indicate one standard deviation.

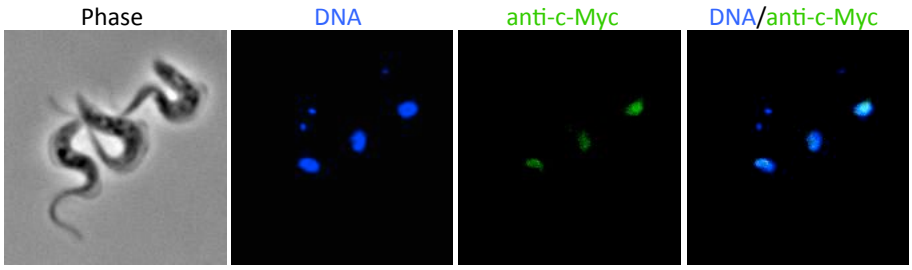


Figure 3.14. TbPriL is a nuclear protein.
A representative *in situ* tagged TbPriL^{Myc} strain was subjected to immunofluoresence analysis with an anti-c-Myc antibody (green) and DNA was counterstained with DAPI. Images were taken with a 100x objective.

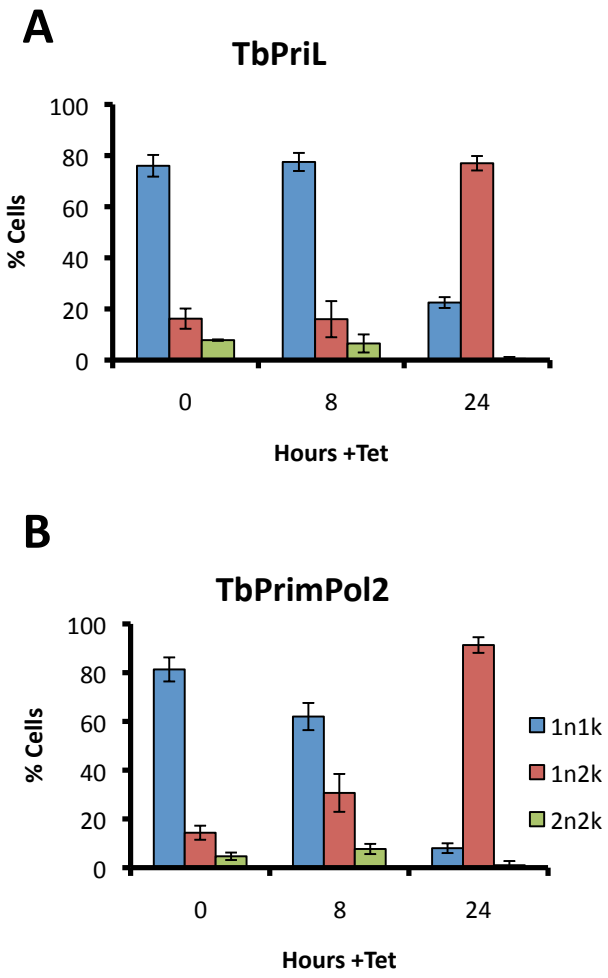


Figure 3.15. TbPrimPol2 and TbPriL are required for cell cycle progression.
TbPriL (**A**) and TbPrimPol2 (**B**) inducible RNAi strains were grown in the presence of 1 µg/ml tetracycline for the times indicated, fixed, stained with DAPI, and analysed by fluorescence microscopy. The ratio of nuclear (n) and mitochondrial kinetoplast (k) genomes was scored per cell (see Figure 3.11a), and the percentage of 1n1k (blue), 1n2k (red), and 2n2k (green) cells calculated. At least 200 cells were scored each experiment, 2 experiments were performed for TbPriL with independent RNAi clones, and 3 experiments for TbPrimPol2 with independent RNAi clones. Error bars indicate standard deviation.

(Figure 3.15b), which is just over one cell division. Such a striking cell cycle arrest following just 24 hours RNAi is unprecedented in bloodstream form *T. brucei*, and strongly supports an essential function of both TbPrimPol2 and TbPriL in cell cycle progression, most likely due to defects during nuclear DNA replication. It is also of note that both TbPriL and TbPrimPol2 knockdown was associated with cells having grossly abnormal morphologies, which is not uncommon for lethal RNAi's, and the nuclei of these cells were also severely distorted (Figure 3.16).

3.8.2. Depletion of TbPrimPol2 arrests cells with 4n DNA content

Accumulation of 1n2k cells is consistent with a defect in nuclear DNA replication. To identify the nature of the replication defect, the DNA content of the cells before and after knockdown was determined using PI staining of DNA coupled with flow cytometry analysis. In uninduced cultures of TbPrimPol2 and TbPriL RNAi strains, the cell cycle distribution was consistent with that of unperturbed cells (Figure 3.17): ~70 % have a DNA content of 2n corresponding to G1, ~15% have 4n DNA content being G2, and the remainder between the two are S-phase cells. Twenty-four hours of TbPriL knockdown prevented cells from fully duplicating their genome and arrested them in S-phase (Figure 3.17). This result in consistent with the established function of the replicative DNA Primase from other eukaryotes and suggests that this proteins essential role is conserved in trypanosomatids. In striking contrast, 24 hours of TbPrimPol2 knockdown resulted in almost all cells stalling with 4n DNA content (Figure 3.17). These data demonstrate that TbPrimPol2 is performing a distinct role from other eukaryotic primases. TbPrimPol2 is not required for initiation or progression of S-phase as knockdown cells can efficiently duplicate the majority, if not all, of their DNA. However, these cells arrest prior to cytokenesis in late-S-G2/M. This phenotype corresponds with the earlier observation that TbPrimPol2 accumulates in G2/M cells (Figure 3.11b), and indicates that TbPrimPol2's essential function occurs after the bulk of DNA synthesis.

3.9. TbPrimPol2 depletion results in the accumulation of DNA damage

The results so far indicate an essential function for TbPrimPol2 that is distinct from previously characterised eukaryotic primases, in that it is required for a process that occurs after the majority of DNA replication takes place. It was next investigated whether the cell cycle arrest in TbPrimPol2 (and TbPriL) knockdown cells coincided with the accumulation of damaged DNA, as this could explain the cell cycle arrest.

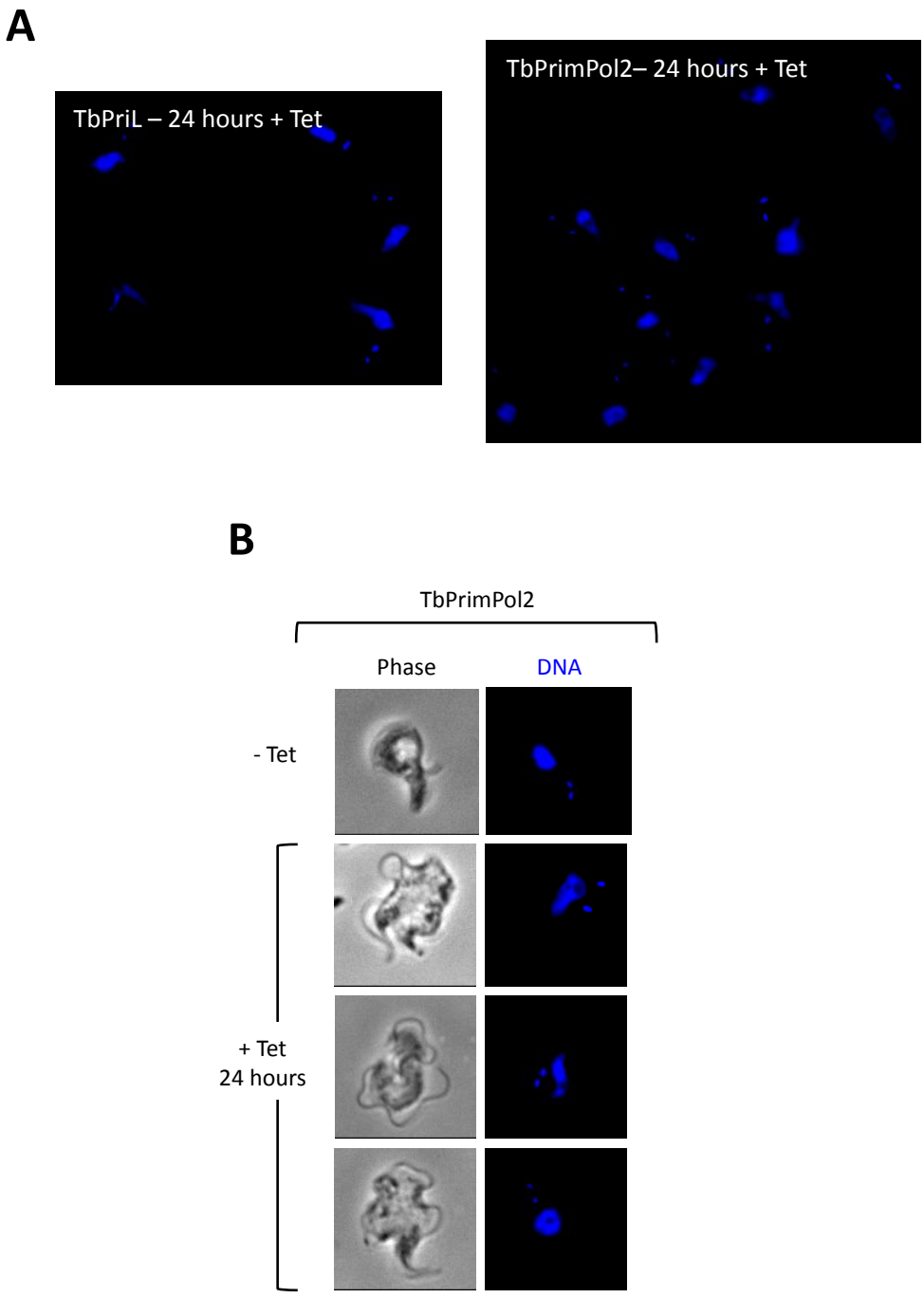


Figure 3.16. Abnormal cell morphology and distorted nuclei observed following TbPriL and TbPrimPol2 knockdown. (**A**) TbPriL and TbPrimPol2 inducible RNAi strains were grown in the presence or absence of 1 µg/ml tetracycline for 24 hours, fixed and stained with DAPI, and analysed by fluoresence microscopy. For TbPrimPol2 knockdown cells, phase contrast and DAPI images are shown (**B**). Images taken with 100x objective.

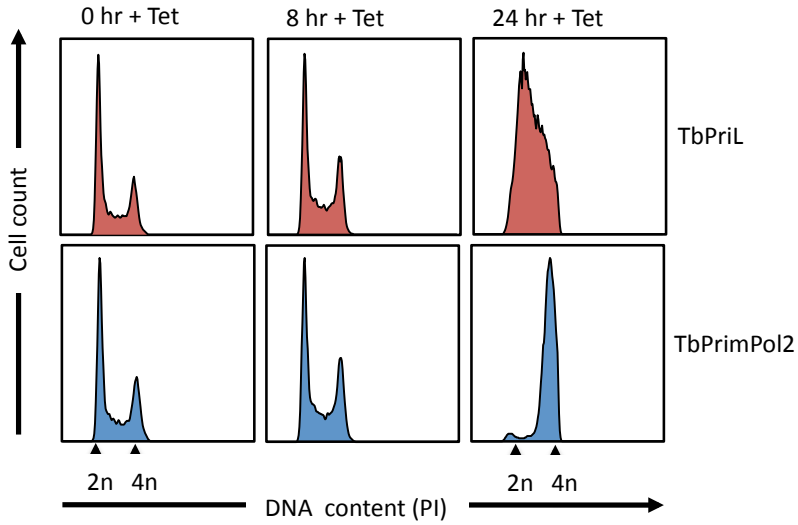


Figure 3.17. TbPrimPol2 knockdown arrests cells with 4n DNA content, whilst TbPriL knockdown arrests cells during S-phase. TbPriL and TbPrimPol2 inducible RNAi strains were grown in the presence or absence of 1 µg/ml tetracycline for the indicated times, fixed and stained with the DNA intercalating dye propiduim iodide (PI), and analysed by flow cytometry. 2n and 4n DNA content are indicated. Representatives shown of experiments performed with 3 independent clones for TbPrimPol2, and 2 independent clones for TbPriL.

3.9.1 TbPrimPol2 depletion causes a small increase in nuclear TUNEL staining

It was first attempted to detect DNA damage using the Terminal deoxynucleotidyl transferase (TdT) dUTP Nick End Labelling (TUNEL) method. TdT catalyses polymerisation of nucleotides to free 3' hydroxyl ends of DNA in a template-independent manner. In the TUNEL assay fixed cells are incubated with TdT and fluorescently labelled nucleotides, and so TdT will fluorescently label free DNA ends that are much more frequent when DNA is damaged. Detection of the fluorophore by fluorescence microscopy allows the visualisation of cells with DNA damage. In an uninduced culture of TbPrimPol2 or TbPriL RNAi strains almost no nuclei were positive for TUNEL staining (Figure 3.18). Twenty-four hours after TbPriL RNAi induction around 2 % of cells contained nuclear TUNEL staining (Figure 3.18). A larger increase was observed in TbPrimPol2 knockdown cells, with around 10 % of cells containing nuclear TUNEL staining (Figure 3.18), suggesting the presence of damaged DNA in these cells. However, this is only a small proportion of cells and is not consistent with the lethality of TbPrimPol2 knockdown.

3.9.2 Depletion of TbPrimPol2 triggers assembly of Rad51 foci

A well-known phenomenon is the assembly of DNA repair and signalling proteins into foci, visible by immunofluorescence microscopy at sites of DNA damage. One such DNA repair enzyme is the recombinase Rad51 that plays a central role in homologous recombination, which repairs DNA double-strand breaks in addition to supporting genome replication (Li and Heyer, 2008). Cultures of TbPrimPol2 and TbPriL inducible RNAi strains were grown in the presence or absence of tetracycline and analysed by immunofluoresence microscopy to detect Rad51. In uninduced cells, a single Rad51 focus was typically detected in ~1 % of nuclei (Figure 3.19), as expected (Proudfoot and McCulloch, 2005; Glover *et al*., 2008). Following 24 hours of TbPrimPol2 knockdown, almost 90 % of cells displayed multiple sub-nuclear Rad51 foci (Figure 3.19). Indeed, almost 15% of cells scored positive for Rad51 foci after only 8 hours TbPrimPol2 knockdown (Figure 3.19). TbPriL knockdown also resulted in the assembly of Rad51 foci, but to a lesser extent. In this case, ~30 % of cells contained multiple sub-nuclear foci after 24 hours of knockdown (Figure 3.19). The accumulation of Rad51 foci following TbPrimPol2 depletion indicates the accumulation of DNA damage in the vast majority of these cells, rather than only in a small proportion as concluded from the TUNEL method. It is possible that the DNA damage caused by either TbPrimPol2 or TbPriL knockdown does not result in free 3' termini of DNA, or at least 3' termini that can be extended by TdT. The accumulation of irreparable DNA damage following TbPrimPol2 knockdown is likely the cause of cell cycle arrest and cell death.

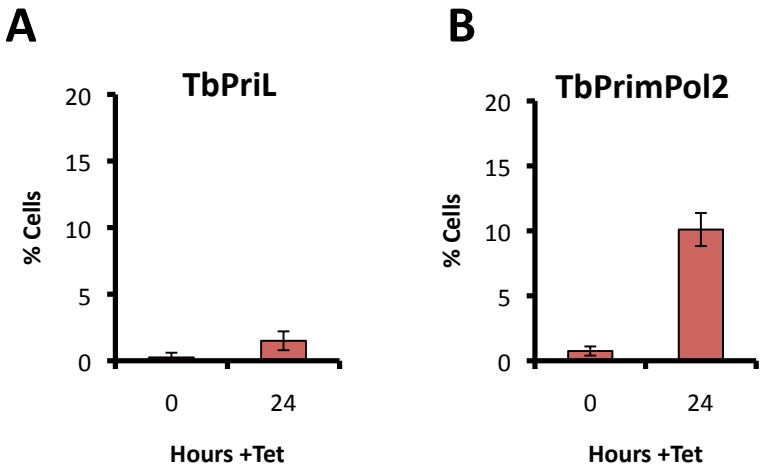


Figure 3.18. TbPrimPol2 RNAi results in a small increase in nuclear TUNEL staining.
Cultures of TbPriL (**A**) and TbPrimPol2 (**B**) inducible RNAi strains were grown in the presence or absence of 1 µg/ml tetracycline for 24 hours and then fixed. Cells were permeabilised and free 3' OH ends of DNA were labelled with fluorescein-labelled nucleotides by TdT for 1 hour (TUNEL method). Cells were then counterstained with DAPI and analysed by fluoresence microscopy and the proportion of cells with nuclear TUNEL staining calculated. Two experiments were performed for each RNAi strain, counting at least 200 cells each time. Error bars represent standard deviation.

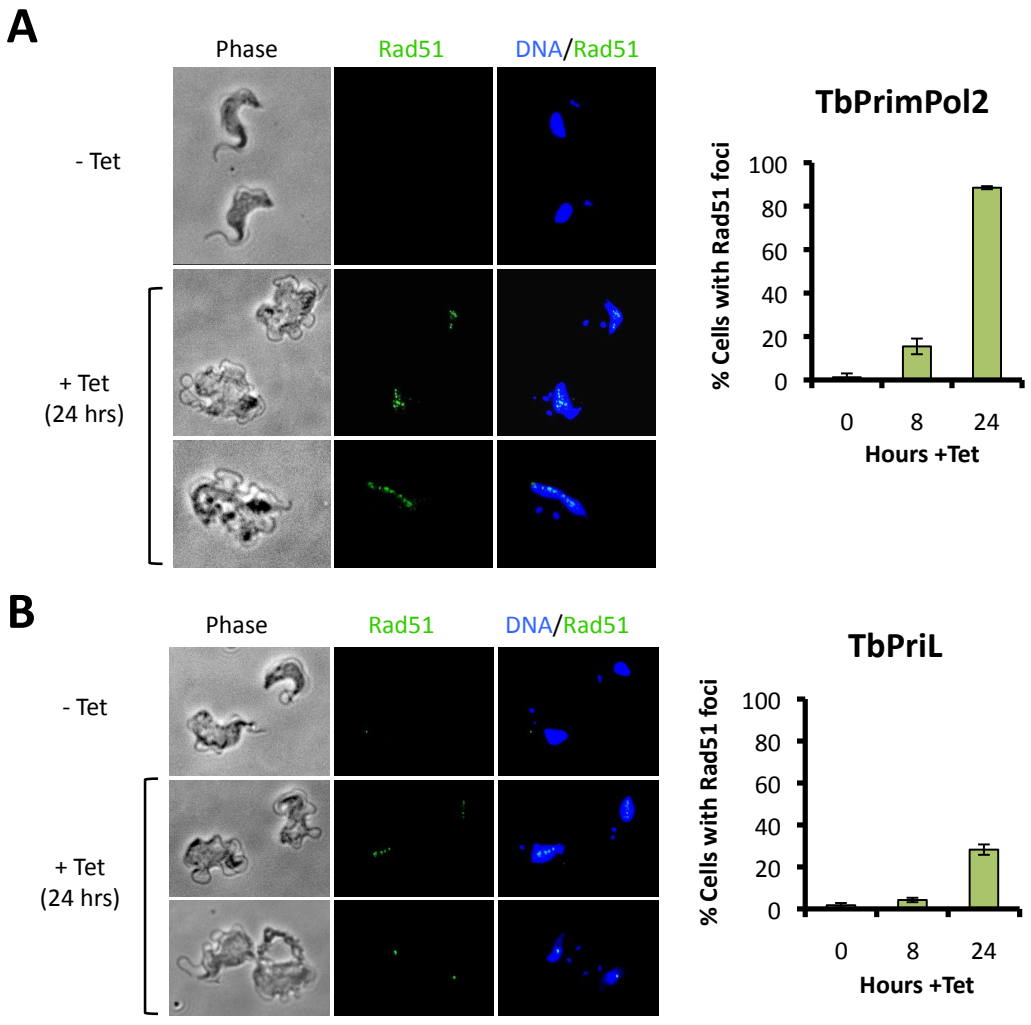


Figure 3.19. TbPrimPol2 and TbPriL depletion triggers assembly of Rad51 foci.
Cultures of TbPrimPol2 (**A**) and TbPriL (**B**) inducible RNAi strains were grown in the presence or absence of 1 µg/ml tetracycline for the times indicated and subjected to immunofluoresence analysis with an anti-Rad51 antibody (green) and DAPI counterstaining (blue). The proportion of cells with sub-nuclear Rad51 foci were counted and representative images of these foci are shown. Two experiments were performed for each RNAi strain, counting at least 200 cells each time. Error bars represent standard deviation

3.10. Over-expression of TbPrimPol2 does not perturb cell proliferation

It was also investigated whether TbPrimPol2 over-expression perturbed cell growth, as was observed for TbPrimPol2 depletion. A *T. brucei* strain was made with inducible over-expression of Myc-tagged TbPrimPol2. The TbPrimPol2 open reading frame without its stop codon was amplified from genomic DNA (see Table 2.4 for primers) and cloned in frame with the coding sequence for 6 consecutive c-Myc epitopes in the vector pRPa^{i6M}, generating construct pRPa^{i6M}:TbPrimPol2 (Table 2.3). Transfection of this construct into the previously described 2T1 strain (see section 3.3) results in the stable integration of the expression construct into the tagged genomic locus, allowing tetracycline-inducible expression of Myc-tagged TbPrimPol2. Two over-expression clones were generated; inducible Myc-tagged TbPrimPol2 (TbPrimPol2^{iMyc}), with a predicted molecular weight of 90 kDa, expressed as a species of approximately 120 kDa following a 24 hour tetracycline induction (Figure 3.20a). Over-expression was not associated with a growth defect after 72 hours (Figure 3.20b).

3.11. Attempt to delete the un-tagged TbPrimPol2 allele in the *in situ* tagged strain

Detection of TbPrimPol2 has thus far relied on an *in situ* Myc-tag at one native allele, however it is not known whether addition of this carboxyl-terminal Myc tag affects TbPrimPol2 function. To determine whether this was the case, it was decided to delete the entire allele encoding for the un-tagged TbPrimPol2, producing a strain that expresses only the Myc-tagged protein. A 527 base pair DNA fragment from the 5' un-translated region (UTR) of the TbPrimPol2 gene, and a 545 base pair fragment from the 3'UTR, were amplified from genomic DNA (see Table 2.4 for primers) and cloned flanking the *PAC* gene (puromycin resistance) in the vector pPAC BS-KS, generating the construct pPAC BS-KS:TbPrimPol2 (Table 2.3) (Figure 3.21a). This construct was transformed into the TbPrimPol2 *in situ* tagged RNAi strain and puromycin clones selected. Five clones were produced and genomic DNA was prepared from these cells and screened by PCR using a forward primer in the 5' UTR and reverse primer in the 3' UTR (Table 2.4). In all clones the wild-type un-tagged PrimPol allele was visible and the Myc-tagged allele was deleted (Figure 3.21b). No further attempts were made to delete the un-tagged TbPrimPol2 allele.

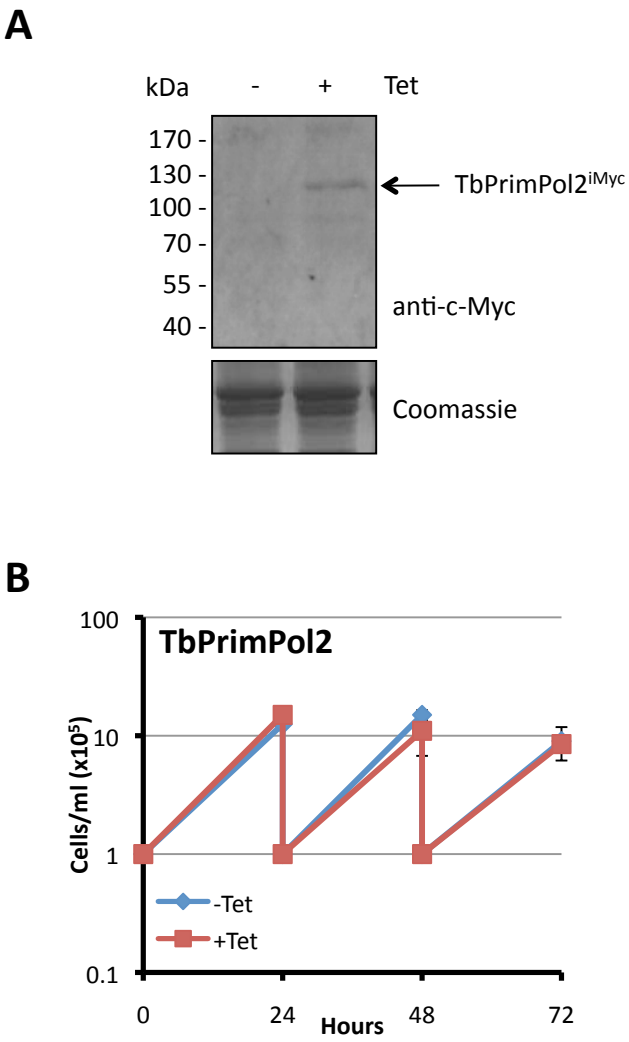


Figure 3.20. TbPrimPol2 over-expression has no effect on cell proliferation.
(**A**) Western blot analysis with anti-cMyc and cell lysates prepared from a representative TbPrimPol2^{iMyc} strain induced for TbPrimPol2 over-expression by addition of tetracycline (1 µg/ml) to culture media for 24 hours. Coomassie stained gel served as a loading control. (**B**) Growth curve of TbPrimPol2^{iMyc} over-expression strain grown in the presence or absence of tetracycline (1 µg/ml). Cultures diluted to 10^5 cells/ml every 24 hours as appropriate. Representative growth curves shown of 2 experiments performed in duplicate, each with independent clones. Error bars represent one standard deviation

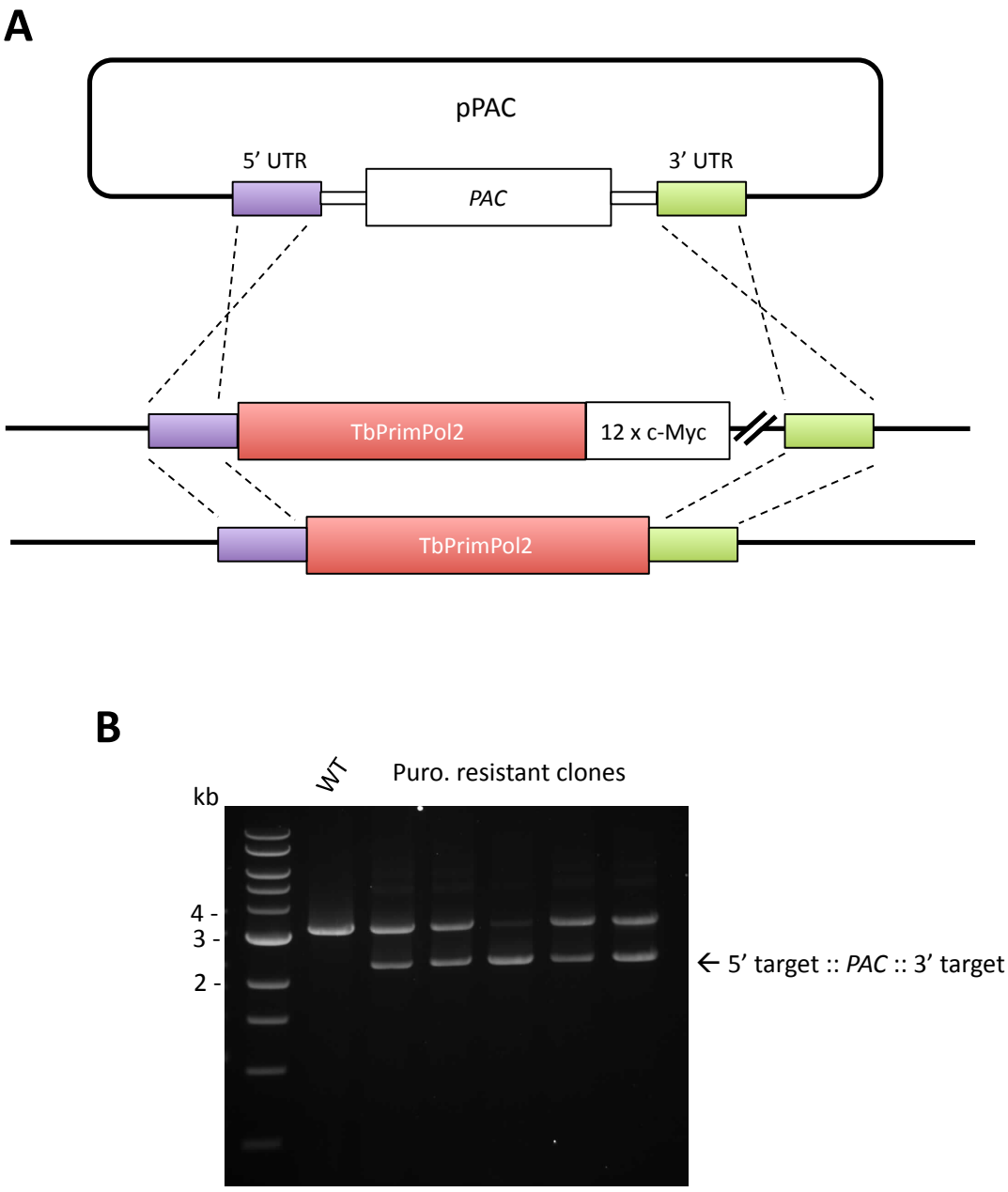


Figure 3.21. An attempt to delete the un-tagged TbPrimPol2 allele in the *in situ* tagged strain. (**A**) Schematic detailing TbPrimPol2 knockout strategy. DNA fragments from the 5′ and 3′ UTR of the TbPrimPol2 gene were amplified from genomic DNA and cloned flaking the puromycin resistance gene (*PAC*). Successful integration of this cassette will result in puromycin resistant clones in which one of the TbPrimPol2 alleles will be deleted. (**B**) Five puromycin resistant clones were generated and screened by PCR. Genomic DNA was prepared from the clones and the presence of the TbPrimPol2 gene tested using PCR with the forward primer used to amplify the 5′ target and the reverse primer used to amplify the 3′ target. The PCR products were resolved by agarose gel electrophoresis alongside the PCR reaction from a wild-type strain.

3.12. Discussion

PrimPol (CCDC111) is a novel eukaryotic primase-polymerase that was originally identified as a member of the AEP superfamily (Iyer *et al*., 2005). The aim of this chapter was to complement the study of higher eukaryotic PrimPol (Chapters 5-7) by characterising PrimPol in a relatively 'simple' eukaryotic model organism. A search of the available eukaryotic genomes identified PrimPol homologues in a wide range of eukaryotes, as previously reported by Iyer *et al.* (2005), but in addition, identified a second group of PrimPol-like proteins in trypanosomatids. Characterisation of these two PrimPol homologues in the important human pathogen *T. brucei* revealed that whilst one of them was dispensable for cell survival, the other was essential and required for a role distinct from previously characterised eukaryotic primases.

Several pieces of data suggest that TbPrimPol2's essential role occurs following the bulk of DNA synthesis in the nucleus. Native allele tagging of TbPrimPol2 revealed this enzyme localised to the nuclei of G2/M cells, in which genome duplication will be advanced, if not complete. RNAi-mediated depletion of TbPrimPol2 arrested cells with a ploidy indistinguishable from 4n with an accumulation of DNA damage, as suggested by assembly of Rad51 foci in almost all cells, and this ultimately led to cell death. These data are consistent with a late- or post-DNA replication role, which is in stark contrast to the role previously described for archaeal and eukaryotic primases, being the synthesis of RNA primers needed to initiate DNA replication. In the nucleus of eukaryotes, this is a role ascribed to the heterodimeric replicative DNA Primase which exists associated to Pol α (Pol α-Prim complex). As would be expected, the putative canonical replicative DNA Primase in *T. brucei* localised to the nucleus and upon RNAi-mediated knockdown, caused cells to irreversibly arrest in S-phase. This presumably was due to the inability of these cells to efficiently initiate DNA synthesis at replication origins and Okazaki fragments, and therefore S-phase progression was severely impeded, consistent with its described role in higher eukaryotic species (Muzi-Falconi *et al*., 2003). The phenotype resulting from knockdown of the replicative DNA Primase is clearly distinct from the phenotype observed following TbPrimPol2 knockdown, as instead of being required for S-phase initiation and progression, this novel enzyme is required for a process that occurs following the majority of DNA synthesis.

Further insights into the role of PrimPol2 in *T. brucei* were gained from work performed in the laboratory of David Horn (LSHTM, London) by Lucy Glover using the strains generated in this chapter. Following knockdown of either TbPrimPol2 or TbPriL almost all cells contained multiple sub-nuclear foci of the phosphorylated histone H2A

(γH2A) (Figure 3.22); a DNA damage signalling protein that is recruited early on at sites of damage, being the trypansomatid equivalent of the higher eukaryotic histone variant γH2AX (Glover and Horn, 2012). This provides further evidence for the accumulation of DNA damage in the absence of TbPrimPol2 (and TbPriL), supporting the observation of Rad51 foci in almost all TbPrimPol2 knockdown cells. It is interesting that despite the apparent abundance of DNA damage following TbPrimPol2 depletion, only a modest level of TUNEL staining was observed, suggesting that the possible DNA damage/lesion does not provide 3' ends compatible with TdT extension. Whatever this DNA damage may be it is clear that it is irreparable, either due to sheer abundance or the nature of the damage, despite the attempted repair of these lesions by homologous recombination, indicated by assembly of Rad51 foci. In addition, further experiments demonstrated that following treatment of cells with the alkylating agent methyl methanesulfonate (MMS), a > 9-fold increase in TbPrimPol2^{Myc} detection was observed, which was a result of TbPrimPol2^{Myc} re-localising into sub-nuclear foci that represented sites of DNA damage, as they co-localised with γH2A (Figure 3.23). This clearly establishes a role for TbPrimPol2 in the DNA damage response. The DNA lesions caused by MMS treatment can stall the replication machinery, and as a result, are substrates for BER and DNA damage tolerance pathways (reviewed in Fu *et al*., 2012). Both of these processes can occur specifically after DNA replication (Otterlei *et al*., 1999; Ulrich, 2011) and so a role for TbPrimPol2 in one of these processes is highly possible, especially in response to endogenously occurring DNA damage, this is investigated further in the following chapter.

Although various pieces of data are indicative of a late- or post-DNA replication TbPrimPol2 function, it is possible that TbPrimPol2 could function during S-phase. As previously detailed, immunfluorescent detection of TbPrimPol2^{Myc} increases following induction of DNA damage (Figure 3.23), and so it is possible that the G2/M detection of TbPrimPol2 in unperturbed cells could reflect these cells containing more DNA damage, being accrued during S-phase, rather than cell cycle regulation of TbPrimPol2 expression. The cell cycle arrest observed in TbPrimPol2 knockdown cells could also be consistent with an S-phase role, with the effects of TbPrimPol2 depletion only becoming apparent following bulk DNA synthesis. However, comparison of the phenotype observed following TbPriL knockdown with TbPrimPol2 knockdown could be informative. Given that TbPriL is a component of the replicative DNA Primase complex involved in S-phase initiation, defects would be expected to arise at the replication fork following TbPriL depletion. Consistent with this, it has been reported in *T. brucei* that problems arising during replication result in the assembly of γH2A foci (Glover and Horn,

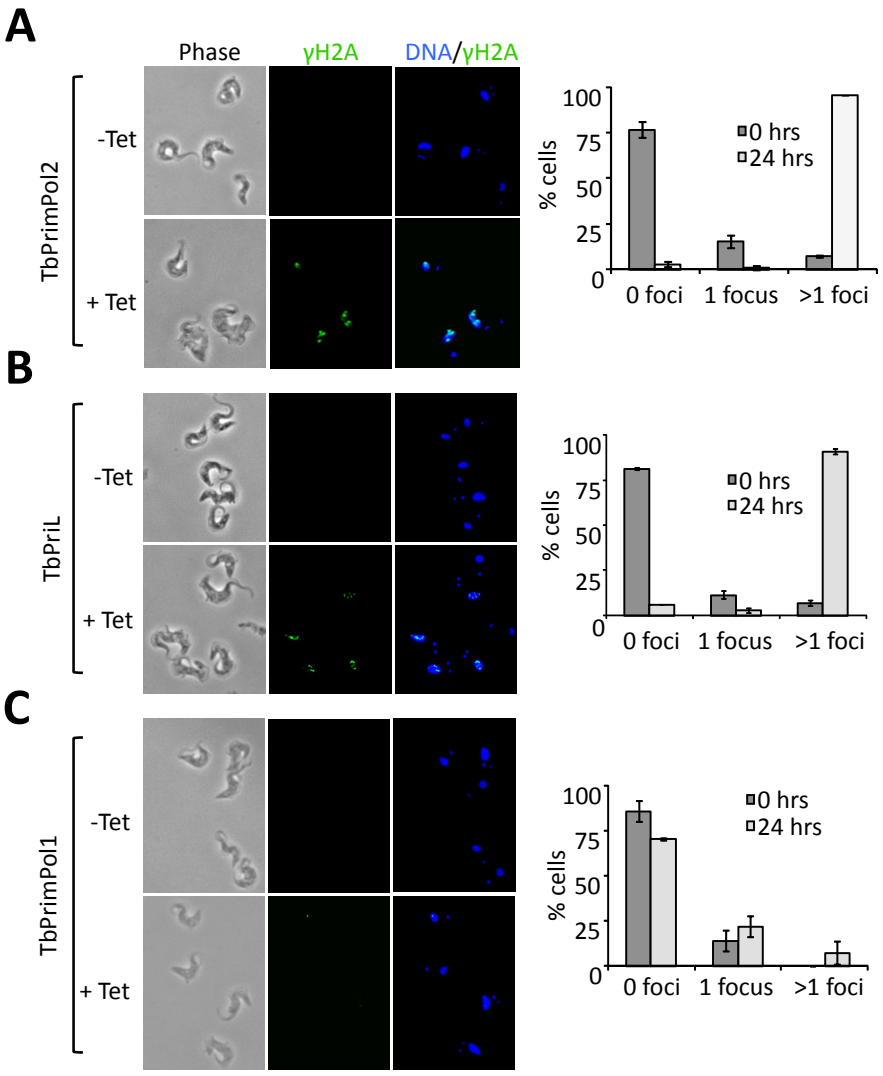


Figure 3.22. Depletion of TbPrimPol2 triggers assembly of γH2A foci.
Immunofluorescent detection of γH2A in (**A**) TbPrimPol2, (**B**) TbPriL, and (**C**) TbPrimPol1 RNAi strains, grown in the presence or absence of tetracycline (1 µg/ml for the time indicated). Representative images including phase-contrast and DAPI stained DNA are shown. The proportion of cells with nuclear foci were counted (n = 200) and error bars represent one standard deviation. Experiment performed by Lucy Glover in the laboratory of David Horn (LSHTM, London).

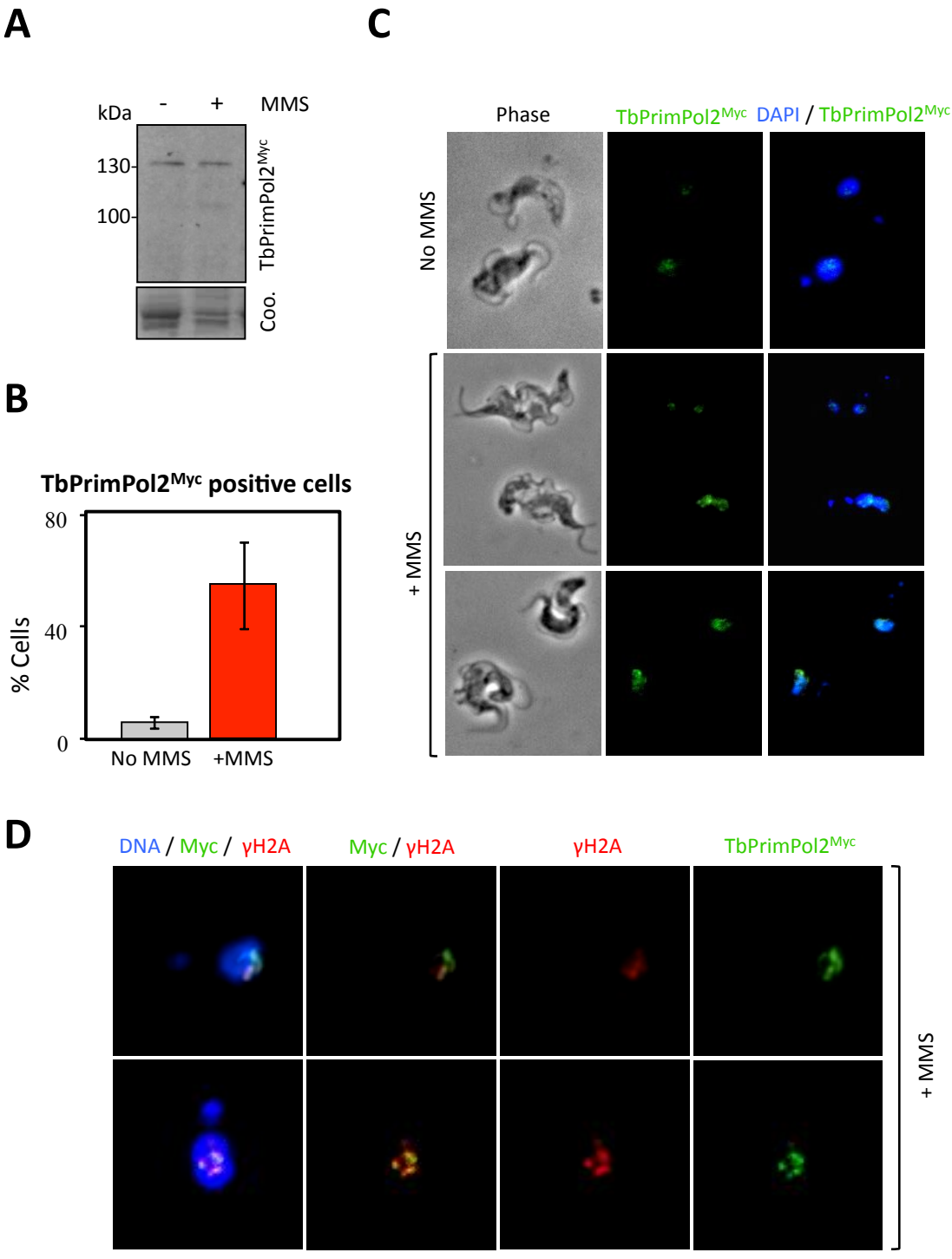


Figure 3.23. TbPrimPol2 accumulates into DNA repair foci.
TbPrimPol2^{Myc} cells were grown in the presence or absence of 0.0003% MMS for 24 hours. (**A**) Western blot analysis with anti-cMyc and cell lysates prepared from MMS treated cells. The Coomassie stained gel serves as a loading control. (**B**) Immunofluorescence analysis of MMS-treated cells with anti-cMyc, the proportion of cells with TbPrimPol2^{Myc} staining was determined. n=200; error bars indicate standard deviation. (**C**) Representative images of cells scored in (B). (**D**) Immunofluorescence analysis of MMS-treated cells with anti-cMyc and anti-γH2A. Experiment performed by Lucy Glover in the laboratory of David Horn (LSHTM, London).

2012) but not Rad51 foci (Glover *et al*., 2008), which is what was observed following TbPriL knockdown (Figure 3.13 and 3.22). In contrast, TbPrimPol2 knockdown results in the vast majority of cells presenting both γH2A (Figure 3.22) and Rad51 foci (Figure 3.13), consistent with a DNA repair defect that occurs downstream of TbPriL function, not at the replication fork. This would be in line TbPrimPol2 functioning after replication fork progression.

The PrimPol homologue that shares marginally more percentage identity with human PrimPol, TbPrimPol1, was dispensable for normal cell proliferation. Attempts to determine TbPrimPol1's sub-cellular localisation proved unsuccessful, perhaps due to low-level expression and/or a dispersed sub-cellular localisation, and, unlike TbPrimPol2, treatment of cells with MMS did not enhance immunofluorescent detection of TbPrimPol1 (Lucy Glover and David Horn, personal communication). However, a small increase in the proportion of cells with sub-nuclear γH2A foci following TbPrimPol1 knockdown was observed (Figure 3.23), suggestive of a nuclear role for this enzyme also. It would be interesting to determine whether TbPrimPol1 (and TbPrimPol2?) had any function in kDNA replication, as it is not uncommon for proteins to be localised to both compartments. But considering that the mitochondria of *T. brucei* already contain two AEP primases, which have been demonstrated to be required for kDNA replication (Hines and Ray, 2010; Hines and Ray, 2011), this may not be the case. Further work would be required to elucidate the role of the dispensable TbPrimPol1.

DNA replication enzymes are potential targets for anti-trypanosomal drugs. The phenotype observed following TbPriL depletion was both rapid and lethal, as was the case for TbPrimPol2 also. Whilst other eukaryotes contain one essential AEP (the replicative DNA Primase), trypanosomes have evolved a DNA replication mechanism that requires two essential AEPs: the replicative DNA Primase to initiate DNA replication and TbPrimPol2, which is essential to complete DNA replication. This indicates that it is not only the mechanism of DNA replication initiation that has diverged in this parasite (Godoy *et al*., 2009; Tiengwe *et al*., 2012a), but a downstream process also. The TbPrimPol2-dependant pathway poses a particularly attractive target for anti-trypanosomal drugs given that this pathway appears specific to trypanosomes, as higher eukaryotic PrimPol is not an essential protein (Julie Bianchi, Laura Bailey, Aidan Doherty, personal communication); further efforts could explore this topic.

Chapter 4

PrimPol1 and 2 in the African trypanosome are trans-lesion synthesis DNA polymerases

4.1. Introduction

Chapter 3 established that a PrimPol homologue in the divergent eukaryote and important human pathogen *T. brucei* was essential for survival. RNAi depletion of this homologue, called TbPrimPol2, caused the accumulation of DNA damage to the nuclear genome that resulted in an irreversible cell cycle arrest following the bulk of DNA synthesis. Further work in the laboratory of David Horn (LSHTM, London) demonstrated that TbPrimPol2 had a role in the DNA damage response, as it re-localised to sites of alkylation DNA damage. It was hypothesised that TbPrimPol2 may function in post-replication repair, a DNA damage tolerance process that occurs after replication fork progression and requires specialised DNA polymerases. The aim of the current chapter was to characterise the *in vitro* activities of TbPrimPol2 (and 1), in particular whether these novel enzymes were capable of TLS, which would be consistent with a role in post-replication repair in the African trypanosome.

4.2. Cloning the *T. brucei* PrimPol1 and PrimPol2 genes

In order to characterise the enzymatic activities of TbPrimPol1 and 2 *in vitro*, the corresponding genes must first be cloned into an appropriate expression vector. The 1827 base pair open reading frame corresponding to TbPrimPol1 (Tb927.5.4070), and the 2199 base pair open reading frame corresponding to TbPrimPol2 (Tb927.10.2520) were amplified from *T. brucei* (Lister 427 strain) genomic DNA (using primers in Table 2.6) and cloned individually into the *E. coli* expression vector pET28a, generating constructs pET28a:TbPrimPol1 and pET28a:TbPrimPol2 (Table 2.5). Each was cloned in frame with an amino-terminal 6-histidine tag and expression of these fusions was under the control of an IPTG-inducible promoter (Figure 4.1a).

4.3. Expression of recombinant TbPrimPol1 and 2 in *E. coli*

The TbPrimPol1 and TbPrimPol2 expression constructs were transformed into B834 (DE3) pLysS (B834s) *E. coli*, the parental strain of the widely used BL21 with the addition of the pLysS plasmid encoding the bifunctional T7 lysozyme protein, which reduces leaky expression in addition to aiding cell lysis (Inouye *et al*., 1973; Moffatt and Studier, 1987; Studier, 1991). Cultures were grown with the addition of zinc to stabilise the putative zinc finger motif of the TbPrimPol's and expression was induced by addition of IPTG. Cells were then lysed and the soluble and insoluble fraction separated by centrifugation before analysis by Western blot with an anti-His antibody. ^{His}TbPrimPol1 has a predicted molecular weight of 70382.3 Da and following addition of IPTG a species of approximately 70 kDa was detected, but was largely in the insoluble cell lysate (Figure

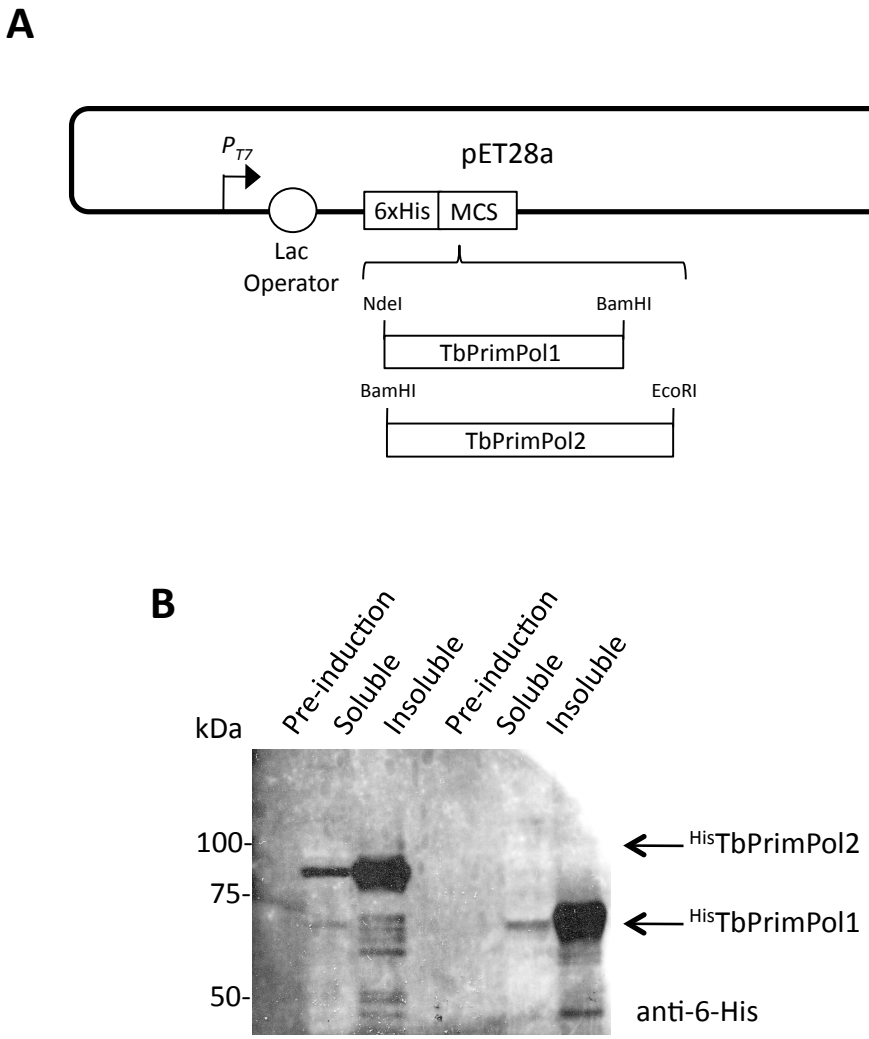


Figure 4.1. Cloning of TbPrimPol1 and 2 and expression in *E. coli*.
(**A**) The open reading frames of PrimPol1 and PrimPol2 were amplified from *T. brucei* genomic DNA (Lister 427) introducing the relevant restriction sites to allow insertion into the multiple cloning site (MCS) of the *E. coli* expression construct pET28a. The genes were cloned in frame with a 6-histidine tag downstream of a T7 promoter (P_{T7}) with a lac operon. (**B**) Cultures of B834s, transformed with either pET28a:TbPrimPol1 or pET28a:TbPrimPol2, were induced with 1 mM IPTG at 30 ^{o}C for 2 hours. Cells were then lysed, sonicated, and the soluble and insoluble fractions separated by centrifugation. A pre-induction whole cell lysate and the induced soluble and insoluble lysate were analysed by Western blot with an anti-6His antibody.

4.1b). ^{His}TbPrimPol2 had a predicted molecular weight of 84146 Da and following addition of IPTG a species of approximately 85 kDa was detected, again largely in the insoluble cell lysate (Figure 4.1b). Although both TbPrimPol1 and 2 expressed well following IPTG induction, further optimisation was required to improve the amount of soluble protein produced.

A number of strategies were attempted to improve the yield of soluble protein; such as the use of *E. coli* strains that express addition tRNAs used in eukaryotes but rarely used in *E. coli* (e.g. Rosetta derivatives), and strains which enhance correct protein folding (e.g. SHuffle). However none proved successful (not shown), and so expression conditions were further optimised in B834s, which have been reported to produce significantly higher amounts of some target proteins when compared to BL21 strains (most strains attempted were BL21 derivatives) (Doherty *et al*., 1995). Various induction times and temperatures were tested, with lower temperatures for shorter times being found to be best, particularly for TbPrimPol2, which started to be degraded after 3 hours of expression (Figure 4.2a). Optimal conditions were a 3 hour induction at 25 ºC, which produced some soluble TbPrimPol1 and 2 (Figure 4.2b). It was decided to purify this soluble protein, at least initially, rather than re-folding the insoluble protein, in order to prevent characterisation of mis-folded protein.

4.4. Purification of recombinant TbPrimPol1 and 2

Cultures of *E. coli* B843s transformed with either the TbPrimPol1 or TbPrimPol2 expression construct were grown until exponential phase and induced for expression by addition of 1 mM IPTG for 3 hours at 25 ºC. The soluble cell lysate was prepared and subjected to Ni^{2+}-NTA agarose affinity chromatography, which will selectively bind the 6-histidine tag fused to both the TbPrimPol's. The bound TbPrimPol's eluted over a range of imidazole concentrations (~70 – 300 mM, not shown), and so to concentrate the elution as much as possible, a single elution of 300 mM imidazole was performed. This successfully eluted TbPrimPol1 and 2, resolving with an apparent molecular mass of ~70 and ~85 kDa respectively on an SDS-polyacrylamide gel (Figure 4.3), which was in agreement with their predicted molecular masses. Both TbPrimPol's co-eluted with a substantial amount of a ~70 kDa *E. coli* contaminant (Figure 4.3a and b) and so further purification was required.

Anion exchange chromatography was first tried to remove the *E. coli* contaminant, which separates proteins on the basis of charge, however both TbPrimPol1 or 2 and the contaminant had the same elution profile (not shown). Rather it was heparin

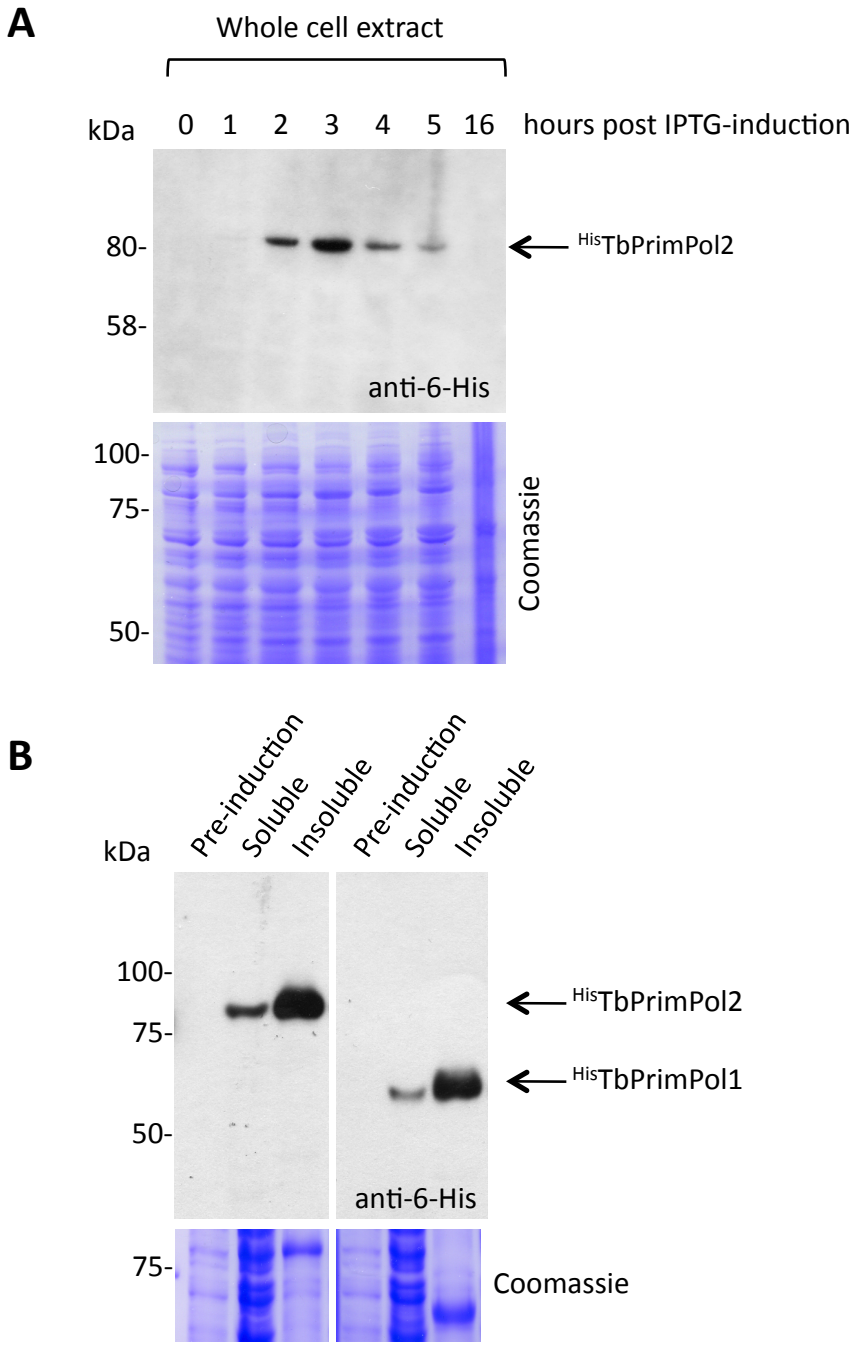


Figure 4.2. Optimising expression of recombinant TbPrimPol1 and 2 in *E. coli*.
(**A**) A culture of *E. coli* strain B834s, transformed with pET28a:TbPrimPol2, was grown to exponential phase (~OD_{600}) and then 1 mM IPTG added to induce ^{His}TbPrimPol2 expression before incubation at 25 ^{o}C. Samples of the culture were taken at the times indicated and the equivalent of 0.15 OD_{600} (~40 µg) was subjected to SDS-PAGE with Coomassie staining and Western blot analysis with an anti-6His antibody. (**B**) Cultures of B834s, transformed with either pET28a:TbPrimPol1 or pET28a:TbPrimPol2, were induced with 1 mM IPTG at 25 ^{o}C for 3 hours. Cells were then lysed and the soluble and insoluble fractions resolved by SDS-PAGE and Coomassie stained, or analysed by Western blot with an anti-6His antibody.

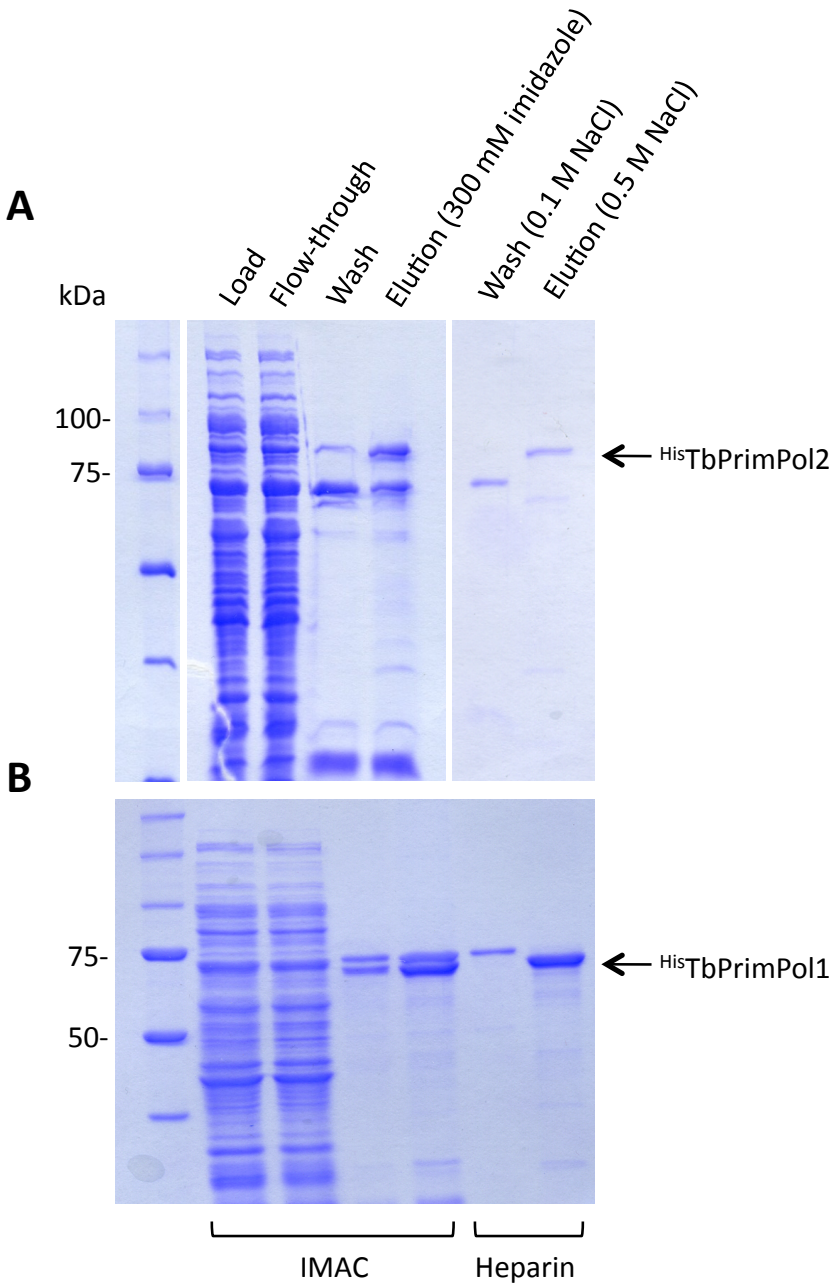


Figure 4.3. Purification of recombinant TbPrimPol1 and 2 from *E. coli*
A 3 L culture of *E. coli* B834s transformed with either pET28a:TbPrimPol2 (**A**) or pET28a:TbPrimPol1 (**B**) was grown to exponential phase and induced by addition of 1 mM IPTG and incubated for 3 hours at 25 ^{o}C. The cells were lysed and the soluble fraction subjected to Ni^{2+}-NTA agarose chromatography. Bound ^{His}TbPrimPol1 and 2 were washed and then eluted with 300 mM imidazole, diluted 10-fold to reduce salt concentration, and then subjected to heparin affinity chromatography, successfully eluting ^{His}TbPrimPol1 and 2 with 0.5 M NaCl. Samples from each purification step were analysed by SDS-PAGE and Coomassie staining, representative gels from ^{His}TbPrimPol1 and 2 purification shown.

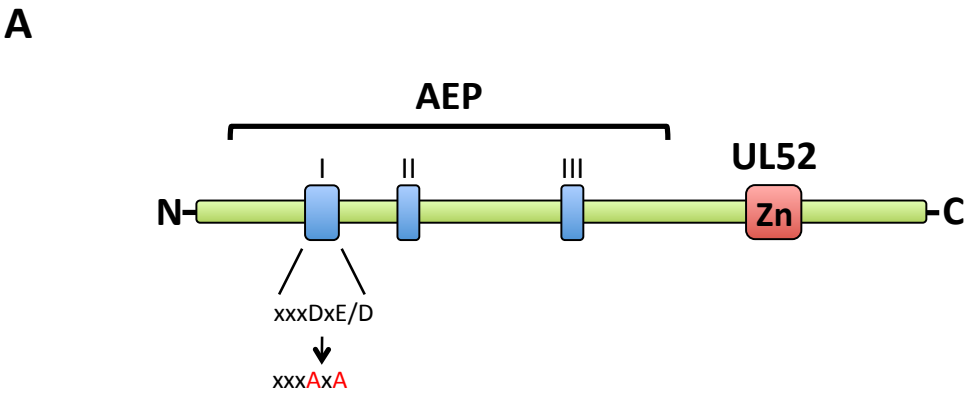
affinity chromatography that succeeded, which selects for DNA binding proteins as it is a phosphate mimetic. The Ni^{2+}-NTA eluate was first diluted at least 10-fold to reduce the salt concentration and then applied to a heparin chromatography column. Both TbPrimPol1 and 2 and the *E. coli* contaminant bound to the column but the contaminant was eluted with a low salt wash, whilst both TbPrimPol1 and 2 eluted at 0.5 M NaCl (Figure 4.3a and b). The typical yield of purified recombinant TbPrimPol1 and 2 was 0.25 mg and 0.1 mg respectively from a 3 litre culture.

4.5. Production of catalytically inactive TbPrimPol1 and 2 mutants

To be certain that the observed enzymatic activities were intrinsic properties of either TbPrimPol1 or 2, catalytically inactive mutants were made. Residues essential for catalysis in TbPrimPol1 and 2 can be predicted based on previous studies of AEP family members (Copeland and Tan, 1995; Augustin *et al*., 2001; Ito *et al*., 2003; Lao-Sirieix and Bell, 2004). The aspartic and glutamic acid residues in motif I of the AEP catalytic domain, predicted to be required for divalent metal binding which is essential to polymerase activity, were selected to be mutated to alanines; D165A E167A in TbPrimPol1 (TbPrimPol1 AxA) and D193A D195A in TbPrimPol2 (TbPrimPol2 AxA) (Figure 4.4). Complementary primers were designed (Table 2.6) to introduce these mutations into the TbPrimPol1 and 2 expression constructs by site-directed mutagenesis PCR, generating constructs pET28a:TbPrimPol1-AxA and pET28a:TbPrimPol2-AxA (Table 2.5). TbPrimPol1 and 2 AxA were expressed and purified in a similar manner to their wild-type counterparts (Figure 4.5).

4.6. TbPrimPol1 and 2 are DNA-dependant DNA polymerases

It was investigated whether TbPrimPol1 and 2 were capable of DNA polymerase activity using the primer extension assay. Primer extensions employ a DNA oligonucleotide primer containing a fluorescent moiety at the 5' end annealed to a longer DNA oligonucleotide template, yielding double-stranded DNA with a 5' over-hang (Figure 4.6). Addition of dNTPs to the 3' end of the labelled primer by a DNA polymerase, in order to replicate the DNA template, will result in labelled products of decreasing electropheretic mobility. Incubation of the primer-template substrate with reaction buffer containing magnesium and either TbPrimPol1 and 2 or dNTPs resulted in no extension of the labelled primer (Figure 4.6, lanes 1, 2, and 7). However, following addition of both, full extension of the primer was observed (Figure 4.6, lanes 3-5 and 8-10). Mutation of active site residues of TbPrimPol1 and 2 resulted in no extension of the labelled primer (Figure 4.6, lanes 6 and 11), indicating the DNA-dependant DNA polymerase activity was an intrinsic property of both TbPrimPol1 and 2. The majority of reaction products



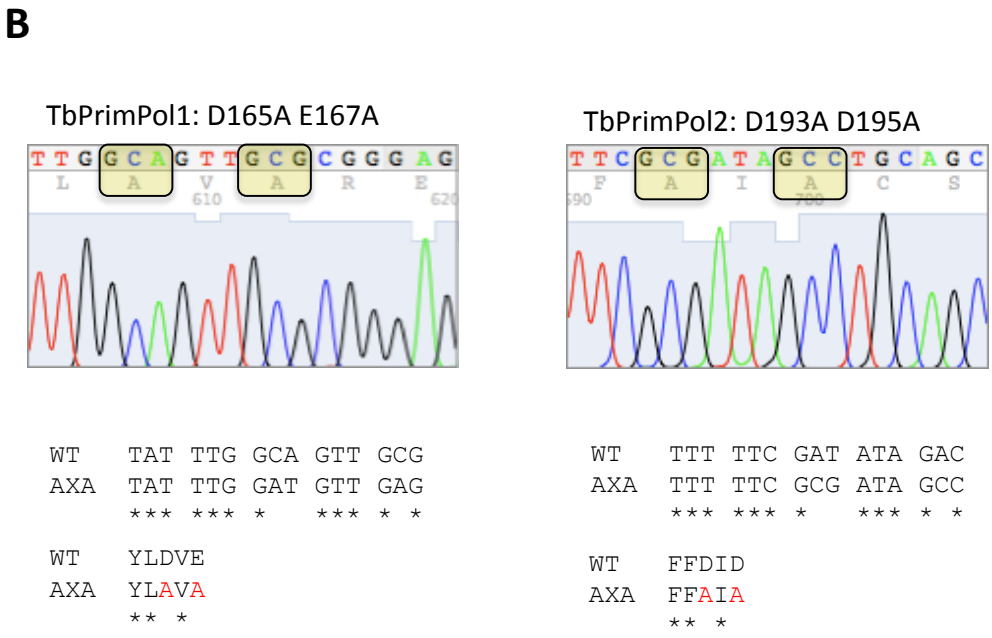


Figure 4.4. Generation of catalytically inactive mutants of TbPrimPol1 and 2.
(**A**) Schematic showing location of residues chosen for site-directed mutagenesis, which are in motif I of the catalytic AEP domain. These residues are predicted to be required for binding of divalent mental ions that are essential for catalysis (**B**) Chromatogram of sequencing reaction from successful site-directed mutagenesis of TbPrimPol1 and 2 expression constructs (TbPrimPol1 AxA and TbPrimPol2 AxA), indicating changes made in nucleotide and amino acid sequence of expression constructs.

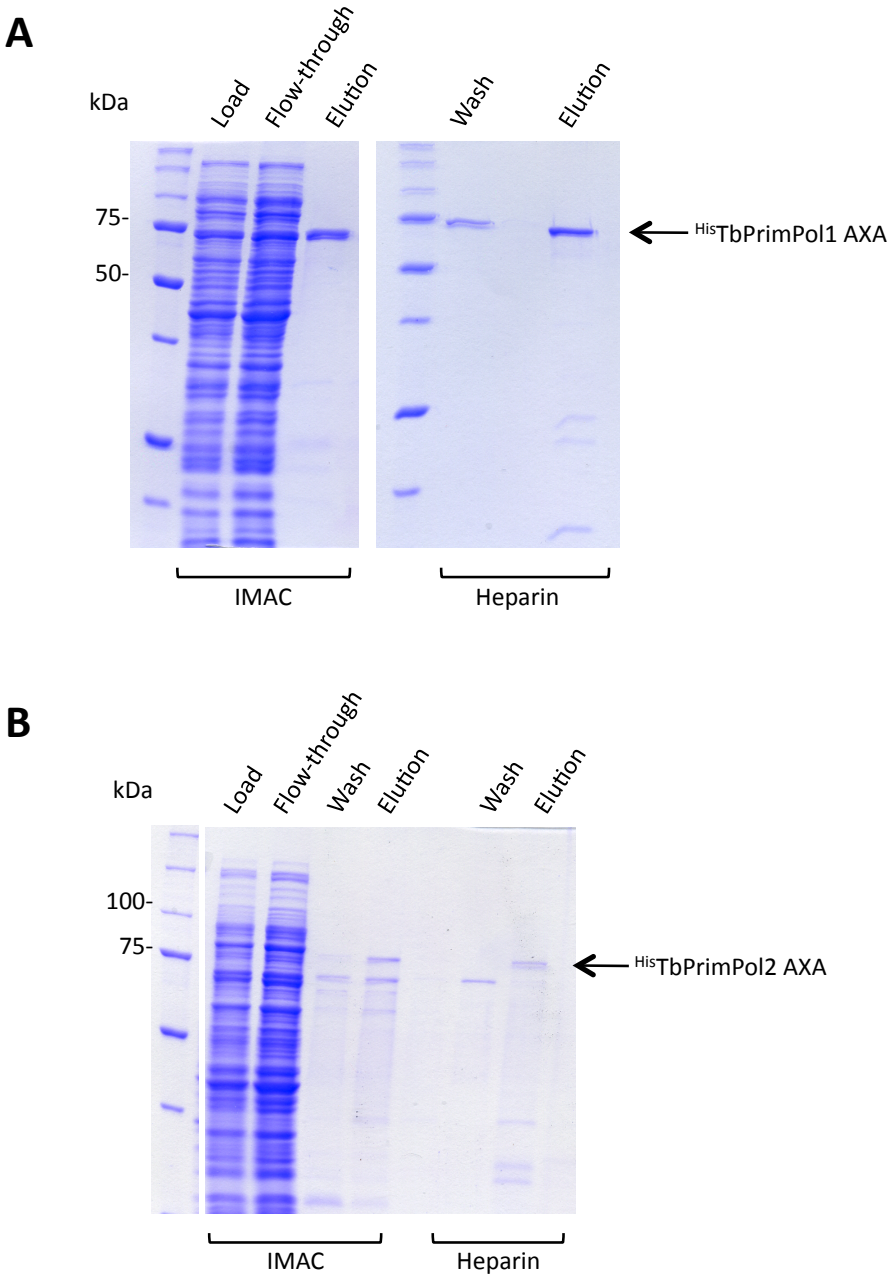


Figure 4.5. Purification of TbPrimPol1 and 2 catalytic mutants.
A 3 L culture of *E. coli* B834s transformed with either pET28a:TbPrimPol1AxA (**A**) or pET28a:TbPrimPol2AxA (**B**) was grown to exponential phase and induced by addition of 1 mM IPTG and incubated for 3 hours at 25 ^{o}C. The cells were lysed and the soluble fraction subjected to Ni^{2+}-NTA agarose chromatography. Bound ^{His}TbPrimPol1 and 2 were washed and then eluted with 300 mM imidazole, diluted 10-fold to reduce salt concentration, and then subjected to heparin affinity chromatography, successfully eluting ^{His}TbPrimPol1AxA and ^{His}TbPrimPol2AxA with 0.5 M NaCl. Samples from each purification step were analysed by SDS-PAGE and Coomassie staining, representative gels from ^{His}TbPrimPol1AxA and ^{His}TbPrimPol2AxA purification shown.

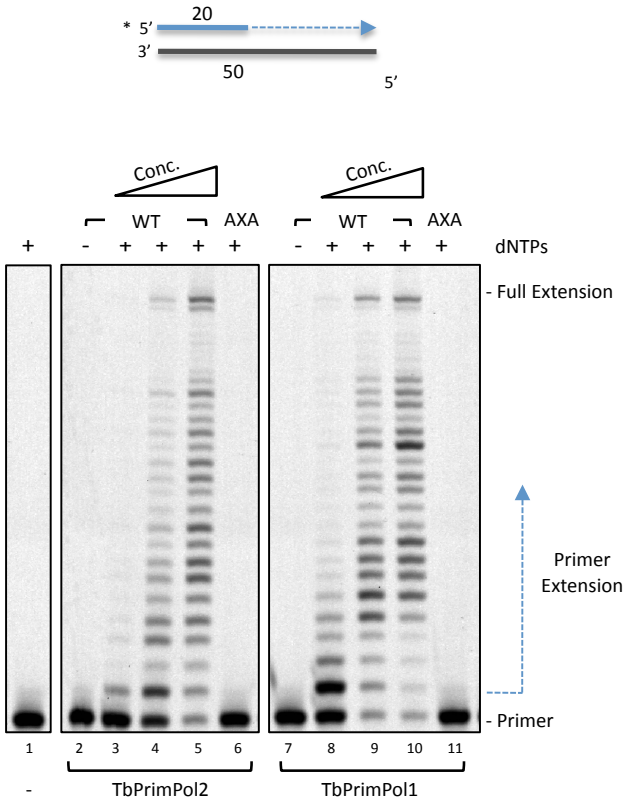


Figure 4.6. TbPrimPol1 and 2 are DNA-dependant DNA polymerases.
The primer-template substrate used in the primer extension assay is shown schematically above the fluorescent scan of the gel; a 50-mer template (black line) annealed to a 20-mer primer (blue line) that contains a 5' fluorescent label (*). Dotted blue arrows indicate the direction of DNA synthesis by a DNA polymerase. Increasing concentrations of recombinant TbPrimPol1 and 2 (50, 125, 250 nM) and the catalytic AxA mutants (250 mM) were incubated with the substrate shown (20 nM), dNTPs (200 µM), and a reaction buffer containing $MgCl_2$, as indicated, for 30 minutes at 37 ^{o}C in a primer extension assay. The reactions were quenched and resolved by DNA-PAGE and the products visualised by detection of the fluorescent moiety of the primer.

synthesised by TbPrimPol1 and 2 were intermediates, even at high enzyme concentrations (Figure 4.6, lanes 5 and 10) suggesting these enzymes are of low processivity, requiring multiple binding events to the primer-template substrate in order to fully extend the primer.

4.7. TbPrimPol1 and 2 are trans-lesion synthesis DNA polymerases

It was next tested whether TbPrimPol2 (and 1) could perform TLS, as this would be consistent with a role for this enzyme in post-replication DNA damage tolerance. The previously described primer extension assay was employed using oligonucleotide templates containing site-specific replication-blocking DNA lesions, so that in order for TbPrimPol1 and 2 to fully extend the primer, they must perform TLS.

4.7.1 Extension opposite a templated T-T CPD

Well-characterised examples of DNA lesions that block replicative DNA polymerases are those caused by UV irradiation, such as CPDs and 6-4 photoproducts (Rastogi *et al*., 2010). A DNA oligonucleotide template was synthesised containing a site-specific *cis-syn* thymine-thymine (T-T) CPD and annealed to a shorter labelled DNA primer for use in a primer extension assay (Figure 4.7a). To be sure that the DNA lesion was a block for replicative DNA polymerases, a primer extension reaction was first performed with a replicative family-B polymerase, using the replicative polymerase from the archaeon *Thermococcus gorgonarius* Tgo-Pol (exo-) (gift from Stanislaw Jozwiakowski, GDSC, Brighton). Consistent with a CPD being a replication blocking lesion (Svoboda and Vos, 1995; Lopes *et al*., 2006), Tgo-Pol was unable to fully extend the primer annealed to the CPD containing template and incorporated a single base opposite the 3' T of the lesion (Figure 4.7a, lane 7). To ensure the integrity of the template beyond the CPD lesion, a modified Tgo-Pol was used that contains the Pol ζ finger domain (Tgo-Z3 Pol) (gift from Stanislaw Jozwiakowski; Jozwiakowski and Connolly, 2011). This polymerase was able to completely replicate the CPD containing template, performing TLS of the lesion, and more importantly ensuring the integrity of the DNA template after the CPD lesion (Figure 4.7a lane 8). TbPrimPol1 and 2 were able to fully extend the labelled primer annealed to the undamaged template (Figure 4.7a, lanes 2 and 3). However, when the primer was annealed to the CPD containing template, both TbPrimPol1 and 2 stalled one base prior to the 3' T of the lesion (Figure 4.7a, lanes 5 and 6). Together, these data indicate that TbPrimPol1 and 2 are incapable of reading through a templated T-T CPD.

TLS can occur by a two-step mechanism, in which one polymerase first incorporates opposite the lesion and a second polymerase extends from this

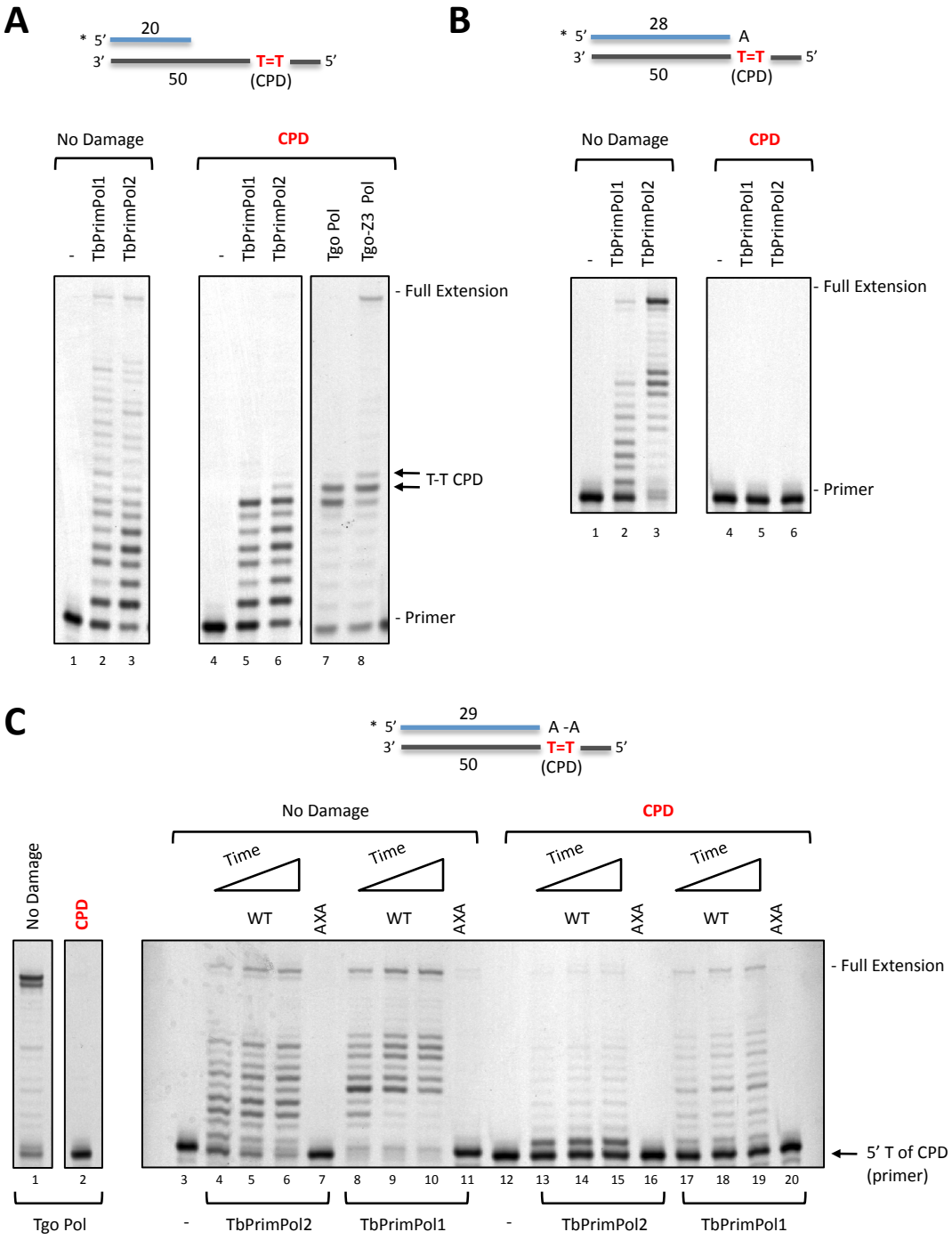


Figure 4.7. TbPrimPol1 and 2 cannot read-through a T-T CPD but can extend from bases opposite. Primer-template substrates used in the primer extension assays are shown schematically above the fluorescent scans of each gel. A 50-mer template (black line) containing a T-T CPD at bases 28 and 29 (numbering from the 3′ side) was annealed to either a 20-mer (**A**), 28-mer (**B**), or a 29-mer (**C**) primer (blue line) with a 5′ fluorescent label (*). The non-damaged template was identical to the damaged except for the absence of a T-T CPD. Reactions consisted of 125 nM of the indicated polymerase, 20 nM substrate, 200 µM dNTPs, an $MgCl_2$ containing reaction buffer, and were incubated at 37 °C for 30 minutes or increasing lengths of time (10, 20, 30 minutes) as indicated. Tgo-Pol and Tgo-Pol-Z3 reactions were for 15 minutes. Following incubation time, reactions were quenched and products resolved by DNA-PAGE and visualised by fluorescent detection of labelled primer.

(mis)matched terminus (Sale *et al*., 2012). A possible extender role for TbPrimPol1 and 2 in CPD bypass was next investigated. A labelled primer was designed and annealed to the CPD template so that a 3' terminal adenine (A) was opposite the 3' T of the dimer (Figure 4.7b). Whilst capable of fully extending from this primer terminus annealed to a non-damaged template (Figure 4.7, lanes 2 and 3), TbPrimPol1 and 2 were completely incapable of extending from this primer terminus when annealed opposite a CPD (Figure 4.7b, lanes 5 and 6), indicating TbPrimPol1 and 2 cannot incorporate opposite the 5' T of a CPD. A second substrate was designed with two 3' terminal adenosines (AA) and annealed opposite both thymines of the templated CPD (Figure 4.7c). The control polymerase Tgo-Pol was completely incapable of extending the primer annealed to the damaged template (Figure 4.7c, lane 2), indicating that this primer terminus when annealed to a CPD lesion is a replication block. TbPrimPol1 and 2 were capable of fully extending the primer annealed to the CPD containing template strand (Figure 4.7, lanes 13-15 and 17-19), however extension was less efficient than on the undamaged template (Figure 4.7 lanes 4-6 and 8-10). The observed TLS activity of TbPrimPol1 and 2 was an intrinsic property of the enzymes as no extension was observed from the catalytically dead mutants (Figure 4.7, lanes 16 and 20). In conclusion, although TbPrimPol1 and 2 could not read-through a T-T CPD, or incorporate opposite the 5' T of the lesion, they could extend from a terminal A-A paired with the CPD lesion, thus performing TLS.

4.7.2. Error-prone bypass of a templated T-T 6-4 photoproduct

The ability of TbPrimPol1 and 2 to replicate DNA containing a 6-4 photoproduct was next investigated. The 6-4 photoproduct is a potent replication-blocking lesion that highly distorts the DNA double helix (Rastogi *et al*., 2010). An oligonucleotide template containing a site-specific T-T 6-4 photoproduct (gift from Alan Lehmann, GDSC, Brighton) was annealed to a shorter labelled primer and used in a primer extension assay. Consistent with a 6-4 photoproduct being a replication blocking lesion, the control polymerase Tgo-Pol was unable to efficiently extend the primer annealed to the DNA lesion containing template, weakly inserting a single nucleotide (Figure 4.8, lane 2). TbPrimPol1 and 2 however, were able to insert nucleotides opposite the T-T 6-4 photoproduct and extend from these nucleotides a total of 6 bases (Figure 4.8, lanes 13-15 and 17-19), performing TLS of a 6-4 photoproduct. This striking activity was intrinsic to TbPrimPol1 and 2 as no extension was observed with the catalytically dead mutants (Figure 4.8, lanes 16 and 20).

As TbPrimPol1 and 2 can read-through a 6-4 photoproduct, it was next investigated whether this bypass was error-free or mutagenic. Primer extension assays

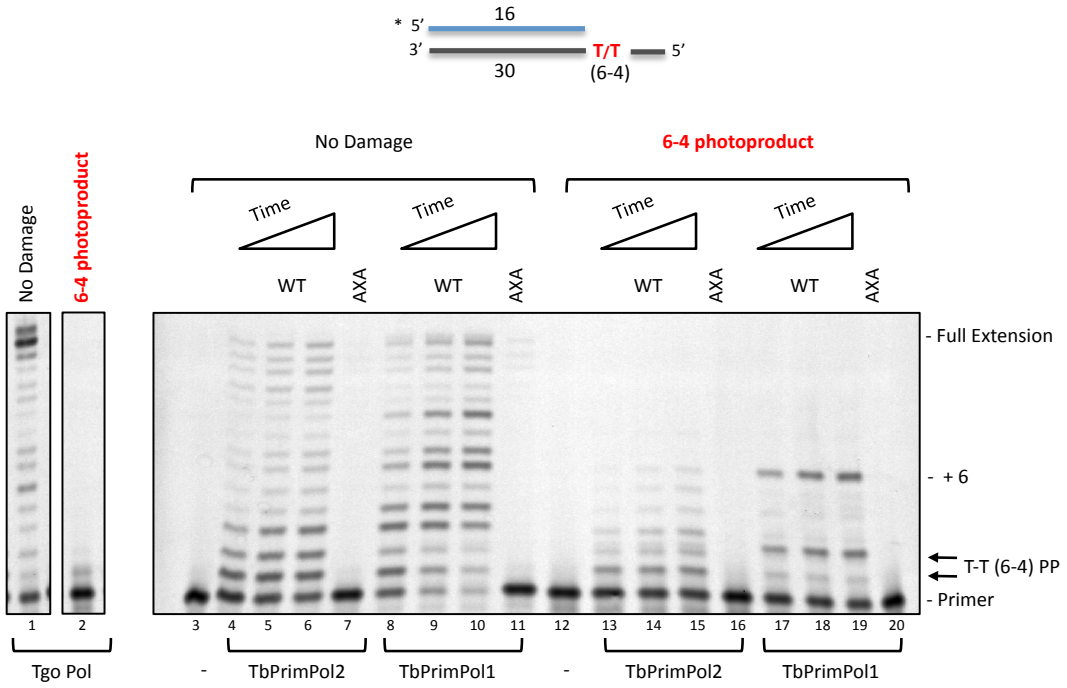


Figure 4.8. TbPrimPol1 and 2 can read-through a T-T 6-4 photoproduct.
The primer-template substrate used in the primer extension assays is shown schematically above the fluorescent scan of the gel. A 30-mer template (black line) containing a T-T 6-4 photoproduct at base 17 and 18 (numbering from 3' side) was annealed to a 16-mer primer (blue line) with a 5' fluorescent label (*). The non-damaged template was identical to the damaged except for the absence of a T-T 6-4 photoproduct. Reactions consisted of 125 nM of the indicated polymerase, 20 nM substrate, 200 µM dNTPs, an $MgCl_2$ containing reaction buffer, and were incubated at 37 ^{o}C. TbPrimPol1 and 2 were incubated for 10, 20, and 30 minutes, AxA for 30 minutes, and Tgo-Pol for 15 minutes before reaction was quenched, resolved by DNA-PAGE, and visualised by fluorescent detection of labelled primer.

were carried out but rather than supplementing the reaction with all four dNTPs, each dNTP was supplied separately. On an undamaged template, both TbPrimPol1 and 2 correctly incorporated two A's opposite an undamaged T-T (Figure 4.9). In the presence of a T-T 6-4 photoproduct TbPrimPol1 and 2 incorporated T opposite the 3' T of the lesion (Figure 4.9). A second substrate was annealed with a 3' terminal T opposite the 3' T of the templated lesion to determine what TbPrimPol1 and 2 insert opposite the 5' T. Both enzymes inserted C or G opposite the 5' T with similar efficiency (Figure 4.9). This choice of incorporation can also be seen in the previous read-through experiments (Figure 4.8, lanes 13-15). In conclusion, TbPrimPol1 or 2-dependant bypass of a T-T 6-4 photoproduct is error-prone.

4.7.3. Incorporate opposite a 3MeA analogue

Replication blocking DNA lesions can also arise from alkylation damage, which can result from both environmental factors and endogenous sources such as metabolic processes (Fu *et al*, 2012). One such DNA lesion is caused by the methylation of adenine, producing 3-methyladenine (3MeA), which accounts for ~20 % of base damage caused by S_N2 methylating agents (Hoffmann, 1980). 3MeA is an inherently unstable lesion and so an oligonucleotide template with a site-specific stable analogue of 3MeA, 3-deaza 3-methyladenine (3dMeA) (gift from Roger Woodgate, NIH, Maryland) (Plosky *et al*., 2008), was annealed to a labelled primer and used in a primer extension assay. Consistent with previous studies demonstrating 3dMeA is a replication blocking lesion (Plosky *et al*., 2008), the control polymerase Tgo-Pol was completely incapable of extending the labelled primer annealed to the 3dMeA containing template (Figure 4.10, lanes 1 and 2). In contrast, both TbPrimPol1 and 2 were able to extend the labelled primer annealed to the 3dMeA-containing template by a single base, incorporating opposite the 3dMeA lesion (Figure 4.10, lanes 13-15 and 17-19). No extension was observed from the catalytically dead mutants (Figure 10, lanes 16 and 20), indicating this TLS activity was an intrinsic property of TbPrimPol1 and 2. It was next investigated whether the TbPrimPol1 or 2-dependant incorporation opposite a 3dMeA lesion was error-free or error-prone using the previously described single incorporation primer extension assays. In both the absence and presence of a templated 3dMeA, TbPrimPol1 and 2 correctly inserted a T, although this insertion was less efficient opposite the damaged base (Figure 4.11). This indicates that TbPrimPol1 and 2 incorporation opposite a 3dMeA is error-free.

4.7.4. Bypass of a templated 8-oxo-guanine

A common source of endogenous DNA damage are free-radical and non-radical oxidants that can modify individual bases, causing lesions such as 8-oxo-G, or cause

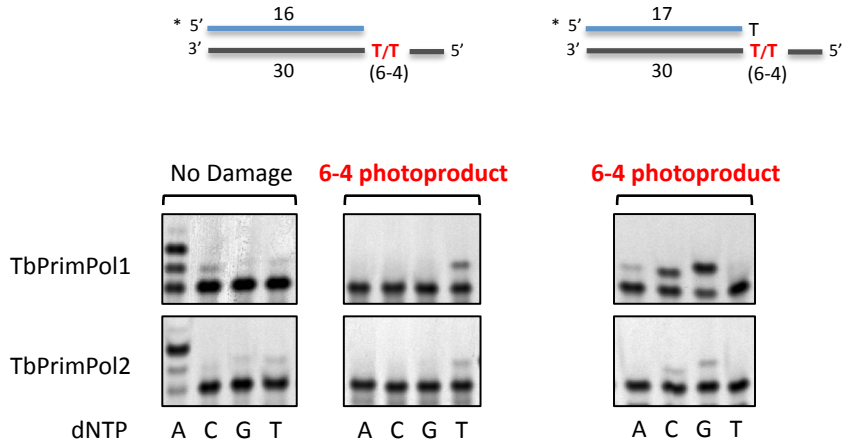


Figure 4.9. TbPrimPol1 and 2 bypass of a 6-4 photoproduct is error-prone.
Primer-template substrates used in the single nucleotide primer extension assays are shown schematically above the fluorescent scans of each gel. A 30-mer template (black line) containing a T-T 6-4 photoproduct at base 17 and 18 (numbering from 3' side) was annealed to either a 16-mer or a 17-mer primer (primer) with a 5' fluorescent label (*). The 17-mer primer had a 3' terminal T annealed opposite the 3' T of the photoproduct. The non-damaged template was identical to the damaged except for the absence of a T-T 6-4 photoproduct. Reactions consisted of 125 nM of the indicated polymerase, 20 nM substrate, 200 µM dNTP, an $MgCl_2$ containing reaction buffer, and were incubated at 37 ^{o}C for 30 minutes. Reactions were then quenched and resolved by DNA-PAGE, and visualised by fluorescent detection of labelled primer.

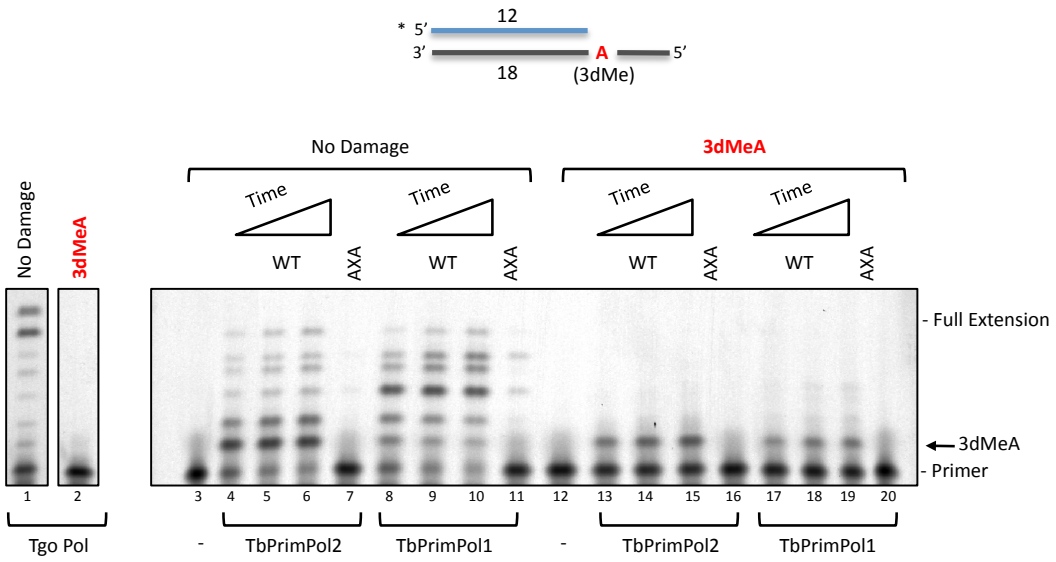


Figure 4.10. TbPrimPol1 and 2 can incorporate opposite a 3MeA analogue.
The primer-template substrate used in the primer extension assays is shown schematically above the fluorescent scan of the gel. An 18-mer template (black line) containing a 3-deaza 3-methyladenine (3dMeA) at base 13 (numbering from 3' side) was annlealed to a 12-mer primer (blue line) with a 5' fluorescent label (*). The non-damaged template was identical to the damaged except for the absence of 3dMeA. Reactions consisted of 125 nM of the indicated polymerase, 20 nM substrate, 200 µM dNTPs, an $MgCl_2$ containing reaction buffer, and were incubated at 37 ^{o}C. TbPrimPol1 and 2 were incubated for 10, 20, and 30 minutes, AxA for 30 minutes, and Tgo-Pol for 15 minutes before reaction was quenched, resolved by DNA-PAGE, and visualised by fluorescent detection of labelled primer.

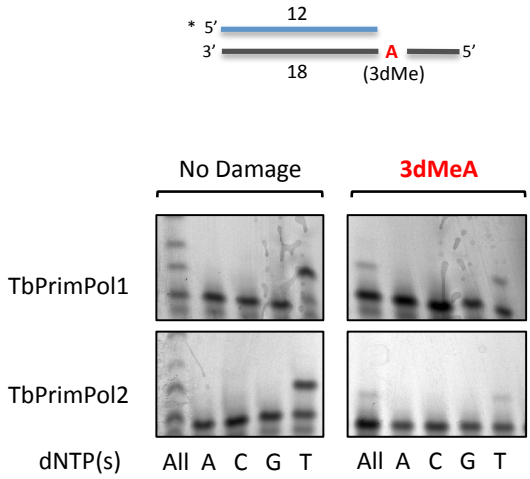


Figure 4.11. TbPrimPol1 and 2 can incorporate the correct base opposite a 3-methyladenine analogue. Primer-template substrate used in the single nucleotide primer extension assays is shown schematically above the fluorescent scans of the gel. An 18-mer template (black line) containing a 3-deaza 3-methyladenine (3dMeA) at base 13 (numbering from 3' side) was annealed to a 12-mer primer (blue line) with a 5' fluorescent label (*). The non-damaged template was identical to the damaged except for the absence of 3dMeA. Reactions consisted of 125 nM of the indicated polymerase, 20 nM substrate, 200 µM dNTP(s), an $MgCl_2$ containing reaction buffer, and were incubated at 37 ^{o}C for 30 minutes. Reactions were then quenched and resolved by DNA-PAGE, and visualised by fluorescent detection of labelled primer.

damage to the sugar backbone causing abasic sites. Although a templated 8-oxo-G does not majorly distort the DNA helix (McAuley-Hecht *et al*., 1994, Lipscombe *et al*., 1995) it can result in stalling or decreased processivity of replicative polymerases, and it has been suggested that a polymerase switch mechanism, such as exists for UV photoproducts, may also exist for 8-oxo-G (McCulloch *et al*., 2009). In yeast, bypass of 8-oxo-G has been suggested to occur in a post-replicative manner (de Padula *et al*., 2004). 8-oxo-G is a mutagenic lesion, as replicative polymerases readily mis-incorporate A opposite a templated 8-oxo-G (reviewed in Berquist and Wilson, 2012). An oligonucleotide template containing a site-specific 8-oxo-G was used in a primer extension assay (Figure 4.12). The control polymerase Tgo-Pol was capable of fully extending the labelled primer annealed to the 8-oxo-G containing template, however significant stalling was observed at the lesion (Figure 4.12, lane 2). Both TbPrimPol1 and 2 could fully extend the labelled primer also, however, a small decrease in processivity was observed when the primer was annealed to the 8-oxo-G containing template (Figure 4.12, lanes 13-15 and 17-19). No extension was observed with the TbPrimPol1 and 2 catalytic mutants confirming that the TLS activity observed was an intrinsic property of TbPrimPol1 and 2 (Figure 4.12, lanes 11 and 20). It was next investigated whether TbPrimPol1 and 2-dependant bypass of an 8-oxo-G was error-prone or error-free. A substrate was annealed so that the 3' terminus of the labelled primer was annealed to the base 3' of the 8-oxo-G lesion. Single incorporation primer extension experiments revealed that both TbPrimPol1 and 2 incorporated A and C opposite a templated 8-oxo-G with similar efficiencies (Figure 4.13), suggesting TbPrimPol1 and 2 dependant bypass of a templated 8-oxo-G can be error-prone and error-free.

4.7.5. Stall prior to a templated abasic site

The ability of TbPrimPol1 and 2 to perform TLS of an abasic site was next tested. In addition to arising directly from DNA damage, abasic sites are also a common intermediate in BER, and can often be more harmful that the original base modification. An oligonucleotide with a site-specific abasic site was annealed to a shorter labelled primer and used in a primer extension assay. TbPrimPol1 and 2 were incapable of fully extending the labelled primer annealed to the abasic site-containing template, both stalling one base prior to the DNA lesion (Figure 4.14a), indicating TbPrimPol1 and 2 cannot read-through an abasic lesion. It was next investigated whether TbPrimPol1 and 2 could function as an extender in abasic site bypass. First, a substrate was annealed with a 3' terminal A opposite the abasic lesion and the ability of TbPrimPol1 and 2 to extend the labelled primer was tested. Both enzymes were completely incapable of extending the labelled primer (Figure 4.14b). A second substrate was tested with a 3'

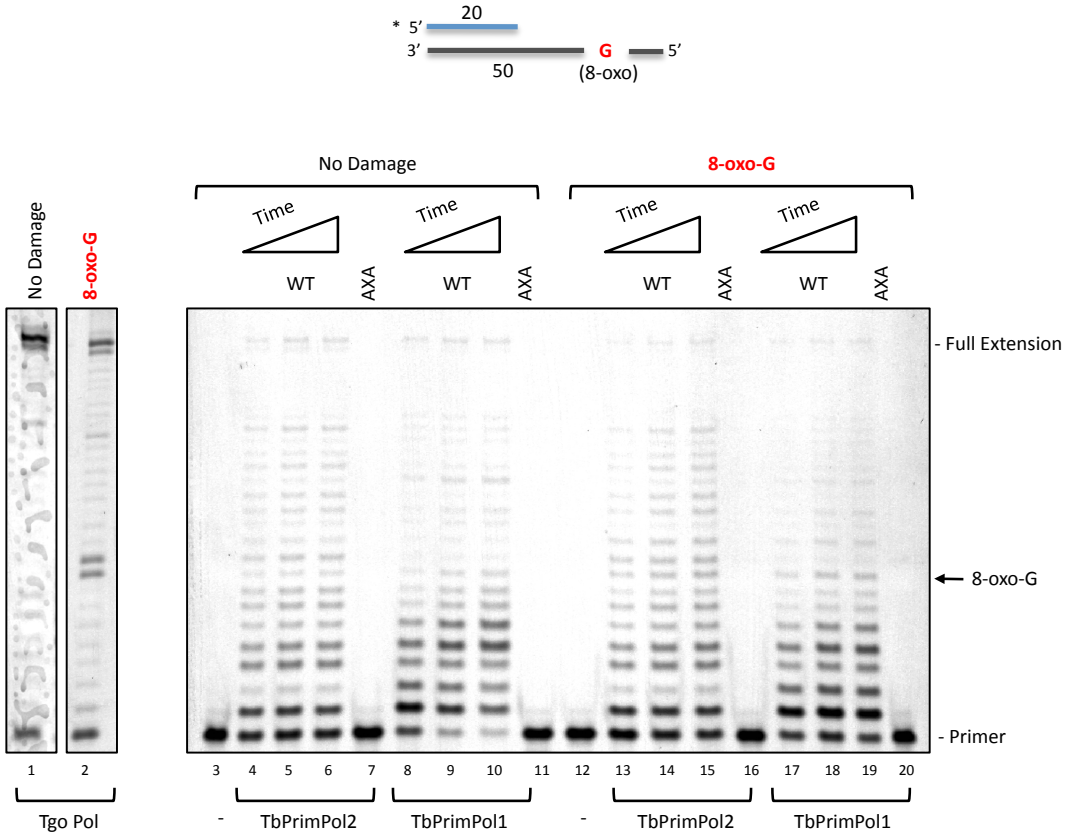


Figure 4.12. TbPrimPol1 and 2 can read-through a templated 8-oxo-Guanine.
The primer-template substrate used in the primer extension assays is shown schematically above the fluorescent scan of the gel. A 50-mer template (black line) containing an 8-oxo-Guanine at base 29 (numbering from 3' side) was annealed to a 20-mer primer (blue line) with a 5' fluorescent label (*). The non-damaged template was identical to the damaged except for the absence of an 8-oxo-Guanine. Reactions consisted of 125nM of the indicated polymerase, 20 nM substrate, 200 µM dNTPs, an $MgCl_2$ containing reaction buffer, and were incubated at 37 ^{o}C. TbPrimPol1 and 2 were incubated for 10, 20, and 30 minutes, AxA for 30 minutes, and Tgo-Pol for 15 minutes before reaction was quenched, resolved by DNA-PAGE, and visualised by fluorescent detection of labelled primer.

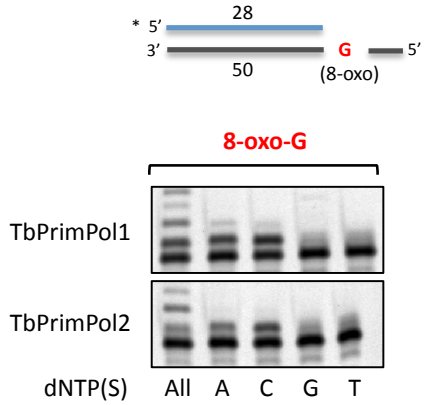


Figure 4.13. TbPrimPol1 and 2 can bypass a template 8-oxo-Guanine in an error-prone and error-free manner. The primer-template substrate used in the single nucleotide primer extension assays is shown schematically above the fluorescent scans of the gel. . A 50-mer template (black line) containing an 8-oxo-Guanine at base 29 (numbering from 3' side) was annealed to a 28-mer primer (blue line) with a 5' fluorescent label (*). Reactions consisted of 125 nM of the indicated polymerase, 20 nM substrate, 200 µM dNTP(s), an $MgCl_2$ containing reaction buffer, and were incubated at 37 ^{o}C for 30 minutes. Reactions were then quenched and resolved by DNA-PAGE, and visualised by fluorescent detection of labelled primer.

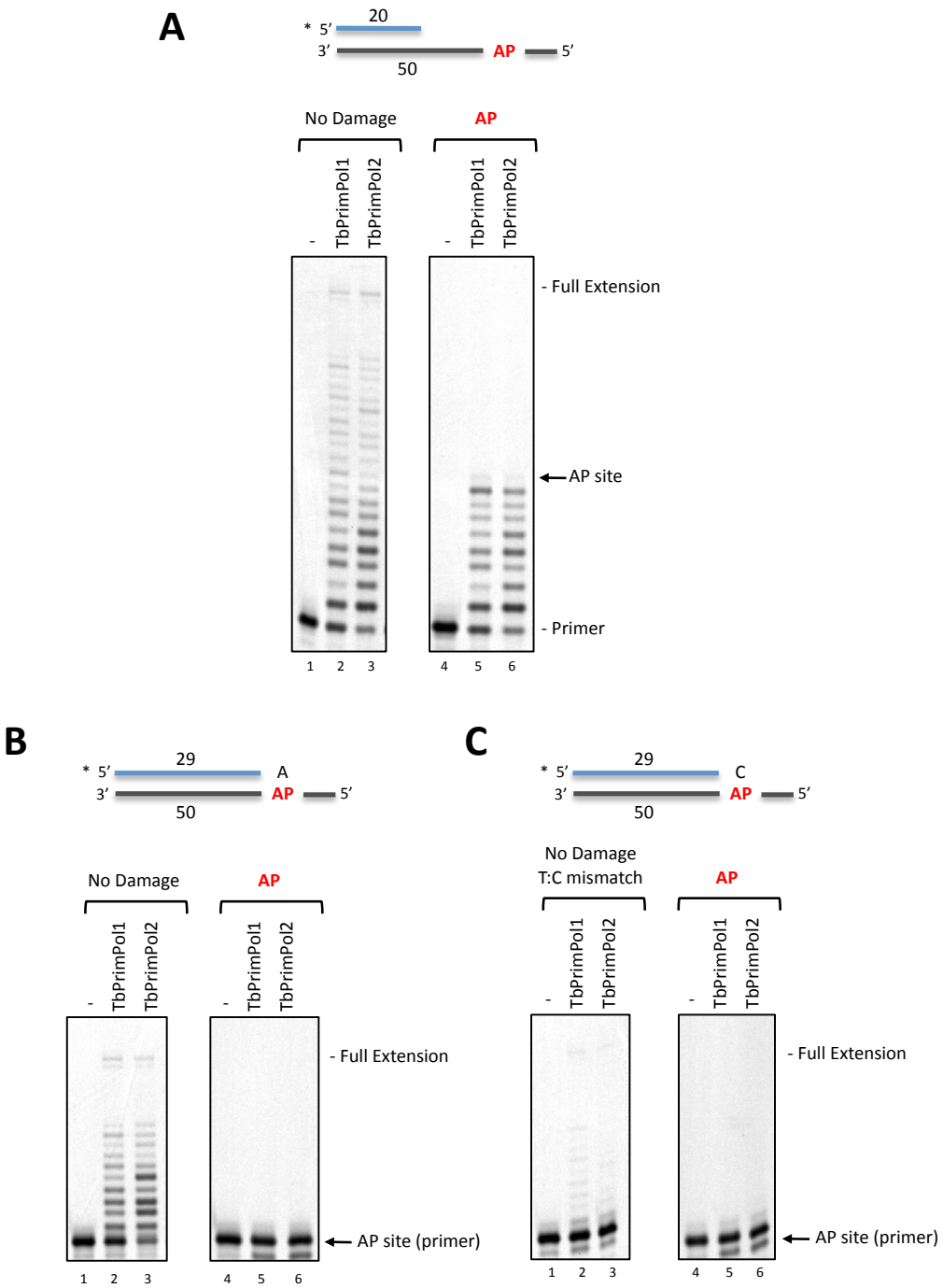


Figure 4.14. TbPrimPol1 and 2 cannot read-through or extend from bases opposite an abasic site. Primer-template substrates used in the primer extension assays are shown schematically above the fluorescent scans of each gel. A 50-mer template (black line) containing an abasic (AP) site at position 29 (numbering from the 3′ side) was annealed to either a 20-mer (**A**), or 29-mer primer (blue line) with a 3′ terminal A (**B**) or C (**C**) annealed opposite the AP site. The primers contained a 5′ fluorescent label (*) and the non-damaged template was identical to the damaged except for the absence of an AP site. Reactions consisted of 125 nM of the indicated polymerase, 20 nM substrate, 200 µM dNTPs, an $MgCl_2$ containing reaction buffer, and were incubated at 37 °C for 30 minutes or increasing lengths of time (10, 20, 30 minutes) as indicated. Tgo-Pol and Tgo-Pol-Z3 reactions were for 15 minutes. Following incubation time, reactions were quenched and products resolved by DNA-PAGE and visualised by fluorescent detection of labelled primer.

terminal C annealed opposite the abasic site, and again no extension of the labelled primer was observed (Figure 4.14c). Both TbPrimPol1 and 2 did weakly extend the primer annealed to the non-damaged template, extending from a T:C mismatch (Figure 4.14c), although this mismatch has been shown not to cause significant difficulty to DNA polymerases (Johnson and Beese, 2004) In conclusion, TbPrimPol1 and 2 cannot perform TLS of an abasic site *in vitro*.

4.8. Discussion

The data presented in this chapter demonstrates that TbPrimPol1 and 2 are DNA-dependant DNA polymerases *in vitro*, and are capable of TLS of replication blocking lesions (summarised in Table 4.1). Both TbPrimPol1 and 2 were capable of extension from bases opposite a T-T CPD, could remarkably catalyse error-prone read-through of a T-T 6-4 photoproduct, in addition to correctly inserting a single dT opposite a 3dMeA, and reading-through an 8-oxo-G incorporating dA and dC equally. TbPrimPol1 and 2 could not read-through a T-T CPD or read-through and extend from an abasic site. These data suggest that TbPrimPol1 and 2 could function in TLS damage tolerance in the African trypanosome, whether reading-through, extending, or inserting opposite a DNA lesion.

The AEP superfamily is proving to be an extremely adaptable group of polymerases. In archaea and eukaryotes primases are required for the essential role of initiating DNA synthesis, and so it was widely accepted that AEPs are DNA-dependant RNA polymerases specialised in primer synthesis (Frick and Richardson, 2001). This was first questioned following the discovery of AEPs in prokaryotes (Koonin *et al*., 2000; Weller and Doherty, 2001), which use the evolutionarily unrelated DnaG Primase for replication, and further questioned with the discovery of versatile archaeal primase-polymerases (Bocquier *et al*., 2001; Lipps *et al*., 2003). These prokaryotic and archaeal AEPs possess a number of nucleotidyl transferase activities and some have been shown to play distinct roles in DNA metabolism, such as DNA break repair in prokaryotes (Pitcher *et al*., 2007b; Brissett and Doherty, 2009). This diversity of AEPs is now extended into eukaryotes with the demonstration that PrimPol is capable of TLS of potent replication blocking DNA lesions. This activity is most likely attributed to the spacious and flexible AEP active site (Lipps *et al*., 2004; Pitcher *et al*., 2007a), which like Y-family polymerases is capable of accommodating major helix-distorting DNA lesions that the replicative polymerases cannot. It is because of the inherent adaptability of the AEP catalytic centre that these polymerases provide a good evolutionary candidate to overcome a range of DNA metabolic problems.

TbPrimPol2 is expressed in the nuclei of G2/M cells, which will have completed the bulk if not all of DNA synthesis, and RNAi depletion of TbPrimPol2 leads to the accumulation of DNA damage and an irreversible cell cycle arrest following DNA replication (Chapter 3). It was hypothesised that this late- or post-DNA replication role could be post-replication repair, a DNA damage tolerance process that occurs following replication fork progression (Ulrich, 2011; Daigaku, 2012), and TLS activity of TbPrimPol2

Table 4.1. Summary of trans-lesion DNA synthesis activities of TbPrimPol1 and 2.

DNA adduct	Trans-lesion synthesis?	Error prone/free?
CPD	extension from dAdA	-
6-4 PP	read-through	error-prone
3dMeA	insertion	error-free
Abasic site	stall prior	-
8-oxo-G	read-through	both

CPD, cyclobutane pyrimidine dimer; 6-4 PP, pyrimidine 6-4 pyrimidone photoproduct; 3dMeA, 3-deaza 3-methyladenine; 8-oxo-G, 8-oxo-2'-deoxyguanosine; dA, deoxyadenosine

is entirely consistent with this hypothesis. Thus, the following model is proposed for the function of TbPrimPol2 (Figure 4.15). In every cell during each cell cycle, despite proficient DNA repair mechanisms, the replication machinery will inevitably encounter abnormal or damaged DNA templates, as a result of endogenous processes and environmental insults. This will result in stalling of the replication fork and so several mechanisms exist to overcome this problem. The DNA lesion can be bypassed at the fork by TLS or template switching, and also bypassed in a post-replicative manner. Replication initiated from adjacent origins can converge on the stalled fork allowing completion of the bulk of DNA synthesis, and replication can re-initiate downstream of the lesion by Okazaki fragment synthesis on the lagging strand and by re-priming on the leading strand (Heller and Marians, 2006; Lopes *et al*., 2006). This will result in a daughter strand with a single-stranded gap opposite the DNA lesion (post-replication gap). The current data suggests TbPrimPol2 is responsible for filling these gaps using its TLS activity, which it does following the bulk of DNA synthesis in late-S/G2 phase (Figure 4.15), thereby restoring the DNA lesion within the safety of the double helix prior to cell division. The phenomenon of post-replication repair has long been documented in bacteria, yeast, and mammals (Rupp and Howard-Flanders, 1968; Lehmann, 1972; Di Caprio and Cox, 1981). More recently it has been demonstrated elegantly in budding yeast that DNA damage tolerance can be completely separate from DNA replication and still be fully operational (Karras and Jentsch, 2010; Daigaku *et al*., 2010). In fact, in both budding yeast and human cultured cells, there is an actual preference for bypassing some DNA damage following the bulk of DNA synthesis rather than during DNA replication (Callegari and Kelly, 2006; Diamant *et al*., 2012). In support of this, the Y-family polymerase REV1 in budding yeast is mainly expressed in G2/M cells, which is when it is suggested to perform its critical DNA damage bypass function (Waters and Walker, 2006), as could be possible for TbPrimPol2 here. In some instances, damage bypass in G2 has been shown to be more mutagenic than bypass during S-phase (Diamant *et al*., 2012), which may be consistent with the increased mutagenicity observed during late replication (Stamatoyannopoulos *et al*., 2009; Lang and Murray, 2011). It is possible that TbPrimPol2 could contribute to the genetic diversity of *T. brucei* through mutagenic bypass of DNA damage.

According to the model, in cells depleted of TbPrimPol2 the post-replication gaps would remain unfilled (Figure 4.15). Post-replication gaps have been shown to contribute to checkpoint activation leading to a G2/M arrest (Karras and Jentsch, 2010; Daigaku *et al*., 2010; Callegari *et al*., 2010), which could be consistent with the 4n arrest observed in TbPrimPol2 knockdown cells in the previous chapter. Also, post-replication

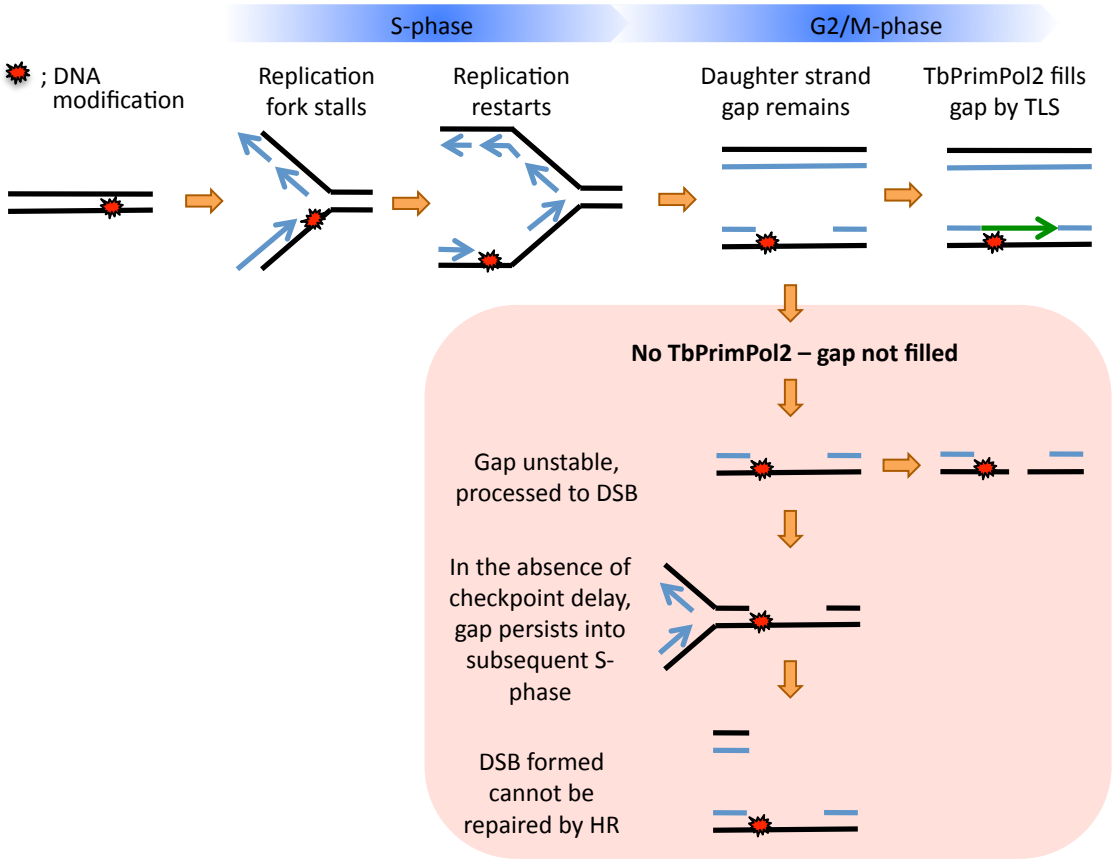


Figure 4.15. Model of the essential role of TbPrimPol2 during cell proliferation.
In all cells during every cell cycle, the DNA replication machinery will inevitably encounter abnormal or damaged DNA templates that result in the replication fork stalling. Replication can proceed despite these damaged templates by replication restart (shown on leading strand) or convergence of adjacent replicons. This will create a single-stranded gap opposite the damaged template. The proposed role of TbPrimPol2 is to fill in these gaps using its trans-lesion DNA synthesis activity (green line), thus restoring the abnormal template in double-stranded DNA and facilitating its later repair. If TbPrimPol2-dependant damage tolerance does not operate, the post-replication gaps will persist and eventually form double-strand breaks in the DNA, this will ultimately lead to cell death.

gaps would be subject to further DNA metabolic processes, lead to double-strand breaks, and be substrates for HR (Elvers *et al*., 2011), which is in line with γH2A and Rad51 foci observed in the absence of TbPrimPol2 (Chapter 3). The phenotype observed in TbPrimPol2 knockdown cells has striking similarity to UV exposed REV1$^{-/-}$ or REV3L$^{-/-}$ mouse embryonic fibroblasts, which are capable of almost completely replicating their DNA, but single-stranded gaps remain encompassing the UV photoproduct that cause these cells to irreversibly arrest in G2 (Jansen *et al*., 2009a and b). One major difference between the studies referenced thus far, and the observations reported here, is that the phenotype resulting from TbPrimPol2 depletion (arrest, DNA damage, cell death) occurs in the absence of external challenges to DNA. This is intriguing as only one other TLS polymerase has been reported to be essential for normal cell proliferation, being mammalian Pol ζ (Lange *et al*., 2012). In the absence of Pol ζ, primary mouse fibroblasts accumulate replication-dependant strand breaks and chromosome aberrations during a single cell cycle, and following further cell divisions ultimately proceed through apoptosis (Lange *et al*., 2012). This again shows similarity to the phenotype observed in TbPrimPol2 depleted cells (Chapter 3), and so it's possible a similar mechanism could be taking place here.

The endogenously occurring DNA damage that is responsible for the severe phenotype observed in TbPrimPol2 depleted cells is unknown. TbPrimPol2 can bypass replication-blocking lesions such as the well-characterised UV photoproducts, but it is very unlikely that UV photoproducts occur naturally in these blood-borne parasites. However, what I have demonstrated is the inherent adaptability of the PrimPol catalytic centre to accept major helix-distorting lesions, such as 6-4 photoproducts, and so it is predictable that PrimPol could bypass other DNA lesions, which should be the topic of further investigation. Common sources of endogenous DNA damage include ROS, which has shown to be a contributing factor to Pol ζ's essential role in mammalian cells (Lange *et al.*, 2012). Also mis-incorporation of ribonucleotides into genomic DNA by the replicative polymerases has also been shown to pose a significant problem to replication (Reijns *et al*., 2012), and it has been demonstrated in budding yeast that a Pol ζ-dependant post-replication repair mechanism is responsible for protecting cells against this (Lazzaro *et al*., 2012), and so TbPrimPol2 may be required for a similar role in *T. brucei*. It is also worth considering that *T. brucei* may be more susceptible to DNA damage; constitutive polycistronic transcription of their genome will involve removal of histones that exposes DNA to possible threats, and the non-transcribed strand will be transiently single-stranded further exposing the DNA to possible threats. Therefore it is conceivable that *T. brucei* may be more reliant on DNA damage tolerance and repair

pathways than other eukaryotes. This may also be the case considering the relatively few replication origins *T. brucei* use to replicate their genome (Tiengwe *et al*., 2012b), as these parasites may be more reliant on damage tolerance to complete genome duplication rather than convergence of adjacent replicons. TLS polymerases are also important for replicating structured DNA and difficult to replicate regions such as chromosomal fragile sites (Bétous *et al*., 2009; Rey *et al*., 2009), and so it is possible that rather than DNA damage, TbPrimPol2 could be required for replication of these difficult to replicate areas of the genome.

This chapter details the cloning, purification, and initial biochemical characterisation of TbPrimPol1 and 2. Further experiments should continue this study and define the catalytic capabilities of these novel enzymes, assaying for activities displayed by other AEPs, such as DNA-dependant RNA/DNA priming, RNA polymerase, and terminal transferase. Additionally, characterising the kinetics and fidelity of these two polymerases would be insightful. Given that TbPrimPol2 is an essential enzyme in bloodstream form *T. brucei*, biochemical and structural studies could develop inhibitors of this enzyme as a tentative drug target, given that higher eukaryotic PrimPol is not an essential enzyme (Julie Bianchi, Laura Bailey, Aidan Doherty, unpublished). Taken together, the work I have presented in this chapter and the preceding chapter describe a novel role for an AEP. Rather than initiation of DNA replication or DNA repair, the data I have presented are consistent with the novel primase-polymerase TbPrimPol2 being required for the tolerance of naturally occurring DNA damage in the pathogenic African trypanosome.

Chapter 5

Localisation of PrimPol in human cells

5.1. Introduction

The human genome is comprised of the DNA in the cell nucleus and mitochondrion. The nuclear genome is approximately 3 billion base pairs in length and exists packaged in 23 chromosome pairs, with only a small fraction of this DNA containing the ~21,000 protein-coding genes (Lander, 2011). In stark contrast, the mitochondrial genome economically compacts 37 genes into a double-stranded circular molecule of just 16,569 base pairs, and is dedicated in its entirety to the aerobic production of ATP (Holt, 2009). Whilst to date 14 DNA-dependant DNA polymerases are known to be located within the nucleus, required for the replication and maintenance of the nuclear genome, only a single DNA polymerase, Pol γ, is present within mitochondria. Pol γ is therefore implicated in all DNA synthetic processes that occur within this essential organelle, being responsible for both the replication and repair of mitochondrial (mt)DNA (Kaguni, 2004). Also in contrast to the nucleus, no DNA primase has been characterised in mitochondria. In the 1980's primase activity from mitochondria was documented (Wong and Clayton, 1985) but the protein responsible was not identified, and it has subsequently been demonstrated that the mitochondrial RNA polymerase (POLRMT) is responsible for priming DNA synthesis (Xu and Clayton, 1986; Wanrooij *et al*., 2008; Fusté *et al*., 2010). Despite the core DNA replication and transcription machineries in mitochondria being distinct from those in the nucleus, many proteins involved in DNA repair are shared between these two organelles (Kazak *et al*., 2012). The human gene *CCDC111* encodes a putative primase, identified by *in silico* analysis as a member of the AEP superfamily (Iyer *et al*., 2005). Unpublished experiments in the laboratories of Aidan Doherty and collaborator Luis Blanco (CBMSO, Madrid) have demonstrated that recombinant human CCDC111 possesses both primase and polymerase activities *in vitro*, and thus the protein was renamed PrimPol. The aim of the work presented in this chapter was to determine PrimPol's sub-cellular localisation, as this would no doubt prove a key component in ascertaining the role of this novel enzyme in human cells.

5.2. Immunofluorescent detection of PrimPol

PrimPol lacks any high probability nuclear or mitochondrial localisation target signals (Aidan Doherty and Luis Blanco, personal communication). Low probability motifs can be identified in a number of prediction programs but they are often located within the predicted catalytic domain and are therefore unlikely to be genuine target signals (not shown). Immunofluorescent microscopy was first used to empirically determine the sub-cellular localisation of human PrimPol. A polyclonal antibody raised against insoluble human PrimPol produced in *E. coli* (gift from Luis Blanco, CBMSO, Madrid) was affinity purified (Julie Bianchi and Aidan Doherty, unpublished) and used to detect endogenous

PrimPol in a number of human cell lines. Initial attempts however generated immunofluorescent images of poor quality (not shown), allowing limited conclusions to be made. Detection of transiently over-expressed PrimPol was next attempted, which was fused to either a green fluorescent protein (GFP) or haemagglutinin (HA) tag (for constructs see Table 2.8), but this again gave inconclusive results (not shown; Julie Bianchi and Aidan Doherty, unpublished). In an attempt to firmly establish the cellular localisation of PrimPol, it was decided to make a cell line stably over-expressing PrimPol fused to an epitope tag.

5.2.1. Generation of a stable, inducible PrimPol expression cell line

To make a cell line stably expressing epitope tagged PrimPol, the Flp Recombinase-mediated Integration (Flp-In) Tetracycline-Regulated Expression (T-REx) system was chosen (overview in Figure 5.1). A derivative of HEK-293 cells (Flp-In T-REx-293) was used that contain a single Flp Recombinase Target (FRT) site integrated into a transcriptionally active genomic locus. The human PrimPol cDNA with a carboxyl-terminal HA tag was amplified from pCI-Neo:PrimPol-HA (primers used in Table 2.9) and sub-cloned into the tetracycline/doxycycline-inducible expression vector pcDNA5/FRT/TO, also containing an FRT site. This generated the construct pcDNA5/FRT/TO:PrimPol-HA (Table 2.8) that was subsequently co-transfected into Flp-In T-REx-293 cells along with a vector expressing Flp recombinase, resulting in the inducible PrimPol expression cassette recombining into the single FRT site. Due to this targeted integration, rather than selecting clones following antibiotic selection, the entire polyclonal population was pooled before testing for expression. A second cell line was also generated expressing PrimPol fused to carboxyl-terminal tandem Flag and *Strep*-Tag-II epitopes, which was mainly used in later chapters. For this cell line the PrimPol cDNA was amplified from pcDNA5/FRT/TO:PrimPol-HA (using primers in Table 2.9) and sub-cloned into pcDNA5/FRT/TO:FLAG-*Strep*-II-Tag (gift from Ian Holt, MRC, Cambridge), generating the construct pcDNA5/FRT/TO:PrimPol-FLAG-*Strep*-II-Tag (Table 2.8). The cell line was generated in the same manner as previously described for HA-tagged PrimPol.

To test for expression of the PrimPol and HA fusion ($PrimPol^{HA}$) cells were grown in the presence or absence of 10 ng/ml doxycycline for 16-24 hours. Western blot analysis with an anti-PrimPol antibody detected endogenous PrimPol in both lysates, but cells grown in the presence of doxycycline also contained a slower migrating species of approximately 70 kDa, consistent with the predicted molecular weight of $PrimPol^{HA}$ (67075.2 Da) (Figure 5.2a). This species was also detected with an anti-HA antibody

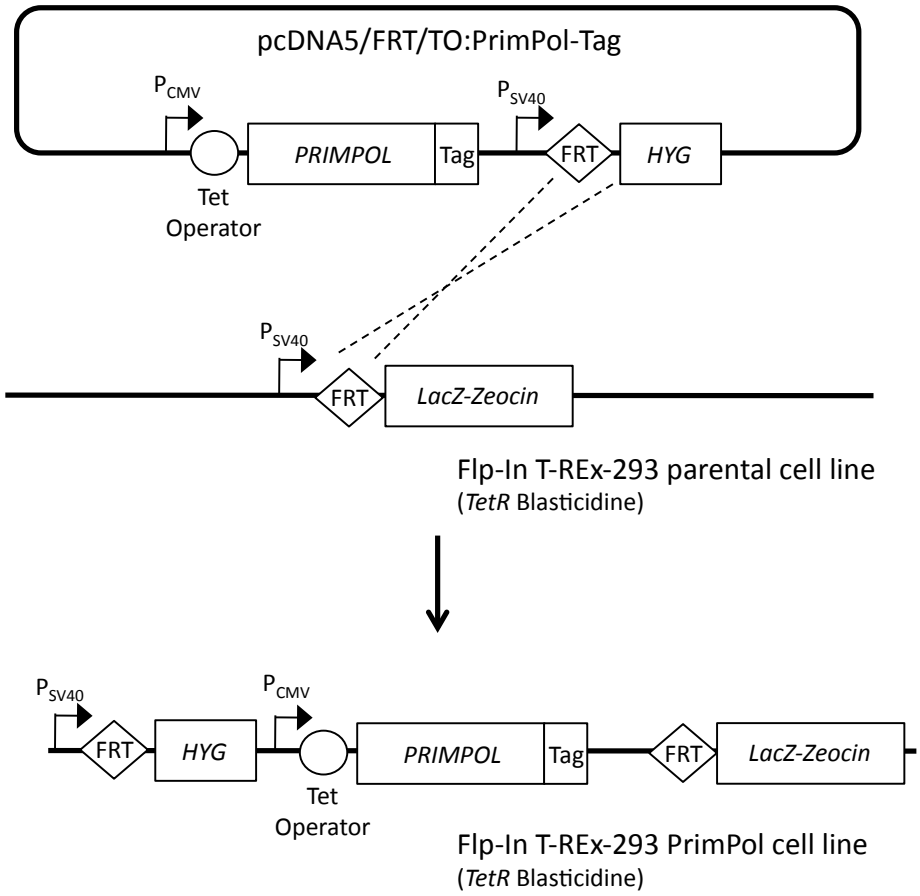


Figure 5.1. Generation of a stable cell line with inducible expression of recombinant PrimPol. The Flp Recombinase mediated Integration (Flp-In) Tetracycline Regulated Expression (T-REx) system was used in HEK-293 (Flp-In T-REx-293) cells. The PrimPol cDNA was sub-cloned into the inducible expression vector pcDNA5/FRT/TO with either a carboxyl-terminal HA or FLAG-*Strep*-Tag-II. The Flp-In T-REx-293 cells constitutively express the Tet repressor (*TetR* Blasticidine) and the LacZ-Zeocin fusion. Correct Flp-recombinase mediated integration at the FRT site replaces the LacZ-Zeocin fusion with the hygromycin resistance gene (HYG), and places expression of PrimPol under a tetracycline/doxycycline-inducible promoter.

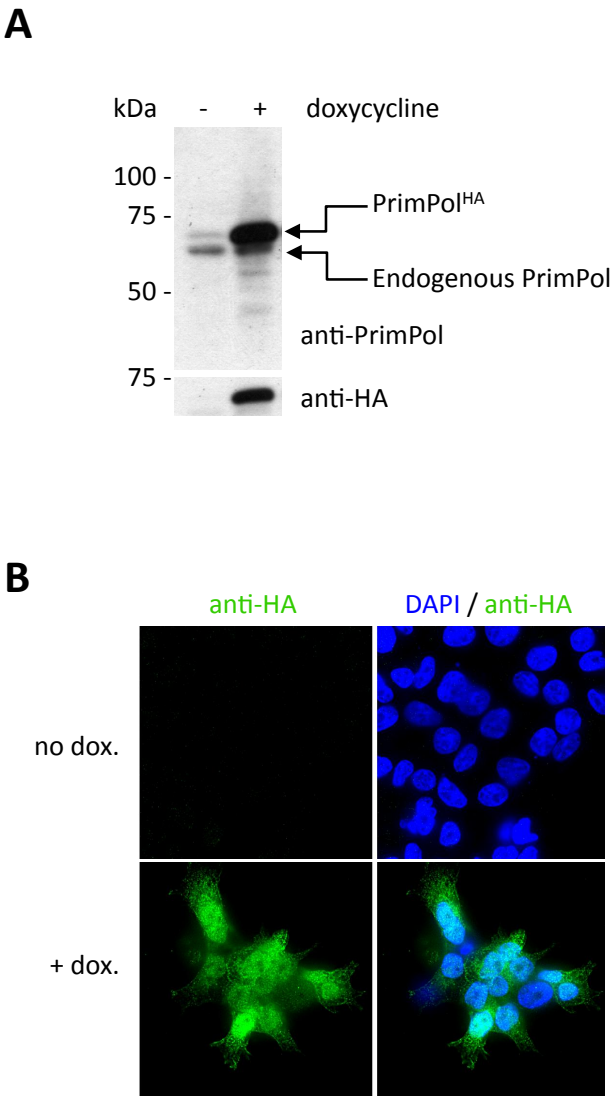


Figure 5.2. Stable, inducible over-expression of recombinant PrimPol in cultured human cells. Flp-In T-REx-293 cells engineered for inducible expression of recombinant PrimPol fused to a C-terminal HA epitope (PrimPol^{HA}) were grown in the absence or presence of 10 ng/ml doxycycline for 16 hours before analysis. (**A**) Cell lysates were prepared and analysed by Western blot with an anti-PrimPol and an anti-HA antibody. (**B**) Fixed cells were subjected to immunofluorescent analysis with an anti-HA antibody (green) and counterstaining with DAPI (blue). Images were taken with a 60x objective and deconvolved.

(Figure 5.2a). Immunofluorescent detection of the HA epitope showed little to no immunofluoresence of cells grown in the absence of doxycycline, however in the presence of doxycycline, a strong signal was observed throughout the cells (Figure 5.2b), consistent with the induction of PrimPol^{HA} expression. Taken together, these data show the successful generation of a stable cell line with inducible expression of HA-tagged PrimPol.

5.2.2. Immunofluorescent detection of stably expressed recombinant PrimPol

PrimPol is capable of primase and DNA polymerase activites and therefore would be expected to localise to the cell compartments containing DNA, being the nucleus and/or the mitochondrion. Cells induced for PrimPol^{HA} expression were counterstained with the DNA intercalating dye DAPI and the mitochondrion-specific dye MitoTracker, to visualise nuclei and mitochondria respectively, and subjected to immunofluoresence analysis with an anti-HA antibody. PrimPol^{HA} immunofluoresence was visible co-localising with both nuclei and mitochondria, however the co-localisation was not complete, with some cytoplasmic immunofluoresence visible (Figure 5.3). Similar results were also obtained with the cell line expressing PrimPol fused to tandem Flag and *Strep*-Tag-II eptitopes (not shown). These data suggest that PrimPol is present in both the nuclei and mitochondria of human cells.

5.2.3. Immunofluorescent detection of endogenous PrimPol

Although recombinant PrimPol over-expressed in human cells localises to nuclei and mitochondria, it is important to determine whether this is the case for endogenous PrimPol. This was difficult however, due to the limitations of the polyclonal antibody available at that time. Soluble human PrimPol was later produced of a high yield and purity using insect cells and this protein was used to affinity purify the polyclonal antibody, and also to generate a new polyclonal antibody (Julie Bianchi, Andrew Green, Aidan Doherty, unpublished), both of which proved much better for immunofluorescent detection of endogenous PrimPol. Various cultured human cells were subjected to immunofluorescent analysis with an anti-PrimPol antibody. Osteosarcoma 143B and U2OS cells, and non-diseased epithelial RPE cells, all gave a punctate PrimPol staining pattern throughout the cell (Figure 5.4). Counterstaining of these cells with DAPI and MitoTracker, to visualise nuclei and mitochondria respectively, revealed partial co-localisation of PrimPol with these organelles (Figure 5.4), but to a lesser extent than over-expressed PrimPol.

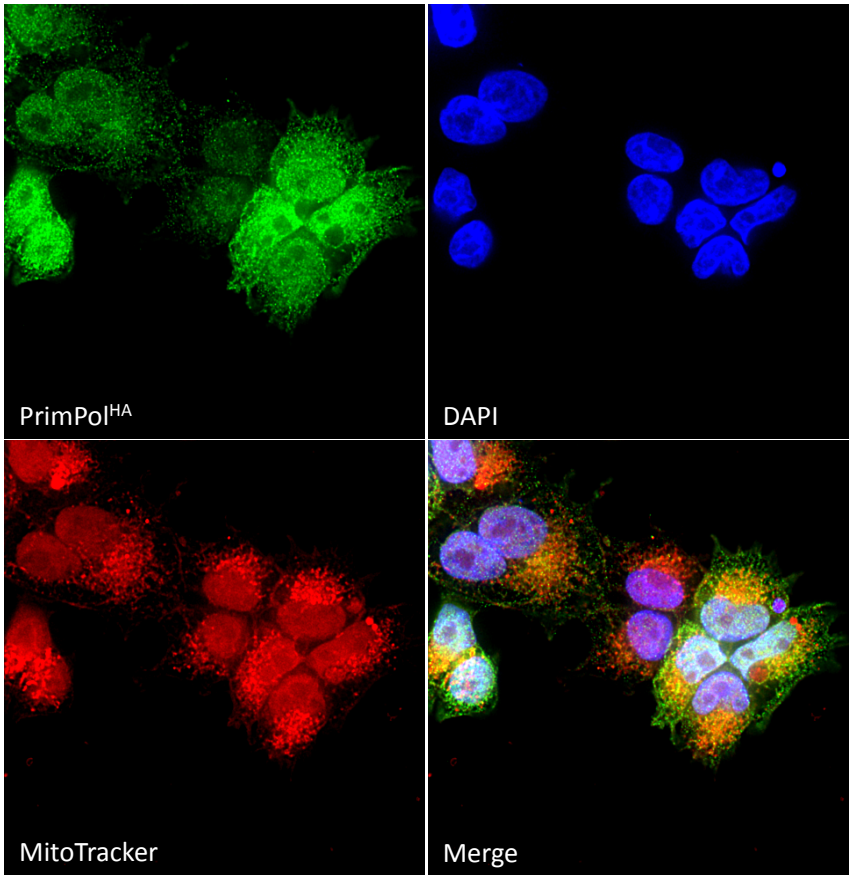


Figure 5.3. Immunofluorescent detection of over-expressed HA-tagged PrimPol in cultured human cells. Flp-In T-REx-293 cells induced for PrimPol^{HA} expression (10 ng/ml doxycycline, 18 hours) were subjected to immunofluorescent analysis with an anti-HA antibody (green). Prior to fixation cells were incubated with MitoTracker Deep Red to stain mitochondria (red), and following immunofluorescent staining were counterstained with DAPI to visualise nuclei (blue). Image was taken with a 60x objective and deconvolved.

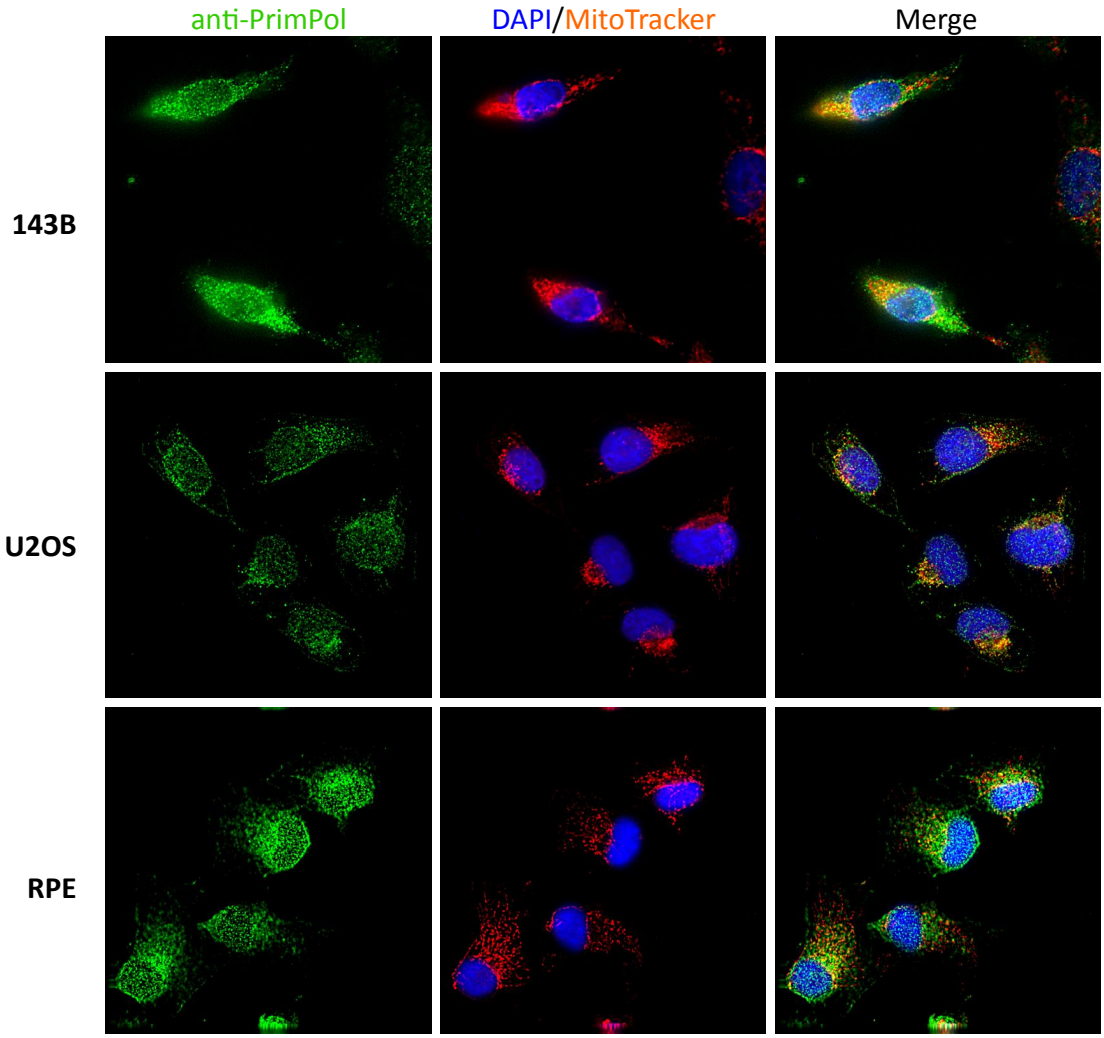


Figure 5.4. Immunofluorescent detection of PrimPol in various cultured human cell lines. Human osteosarcoma 143B and U2OS cells, and non-diseased epithelial RPE cells, were subjected to immunofluorescent analysis with an anti-PrimPol antibody (green). Prior to fixation cells were incubated with MitoTracker Deep Red to stain mitochondria (red), and following immunofluorescent staining cells were counterstained with DAPI to visualise nuclei (blue). Images were taken with a 60x objective and deconvolved.

To determine whether the staining pattern observed was specific to the PrimPol protein, cells were transfected with a PrimPol siRNA prior to immunofluorescent staining. Western blot analysis of lysates prepared from these cells with an anti-PrimPol antibody shows the specific loss of an approximately 65 kDa species corresponding to PrimPol (Figure 5.5a). Immunofluorescent analysis of these cells revealed a substantial reduction of PrimPol immunofluorescent staining (Figure 5.5b), verifying the staining pattern was dependant upon the PrimPol protein. Taken together these data suggest that PrimPol is distributed throughout the cytoplasm, with some protein present in nuclei and mitochondria.

5.3. PrimPol is a mitochondrial protein

Immunofluoresence studies of over-expressed and endogenous PrimPol suggest a portion of this protein is mitochondrial in human cells. To investigate this further sub-cellular fractionation experiments were performed on cultured human cells.

5.3.1. PrimPol is present in a mitochondrial preparation

Intact mitochondria were isolated from HEK-293 cells by differential centrifugation and their contents analysed by Western blot, along with the contents of the cytosol. The mitochondrial DNA polymerase accessory subunit PolG2 was detected exclusively in the mitochondrial preparation (Figure 5.6a), as would be expected, confirming the successful isolation of mitochondria. The detection of the cytoskeletal component β-actin was intended as a cytosolic marker but, consistent with a recent report (Reyes *et al*, 2011), it was also present in the mitochondrial preparation (Figure 5.6a). Although the majority of PrimPol was detected in the cytosolic fraction it was also visible in the mitochondrial preparation (Figure 5.6a). A similar result was also observed with mitochondria prepared from U2OS cells (not shown), which is consistent with the immunofluorescent studies, suggesting a portion of PrimPol is mitochondrial.

5.3.2. PrimPol is retained in protease-treated mitochondria

To confirm the presence of PrimPol inside mitochondria, rather than the possible association with the outer-mitochondrial membrane, a protease protection assay was carried out. Whole mitochondria were isolated from U2OS cells and then treated with the protease Proteinase K in the absence or presence of the detergent Triton X-100, and the remaining proteins analysed by Western blot. The mitochondrial protein PolG2 was detected in the mitochondrial preparation and the level of this protein remained the same following addition of Proteinase K (Figure 5.7, compare lanes 2 and 3 as more protein was loaded in lane 1). Following addition of Triton X-100, which permeabilises the

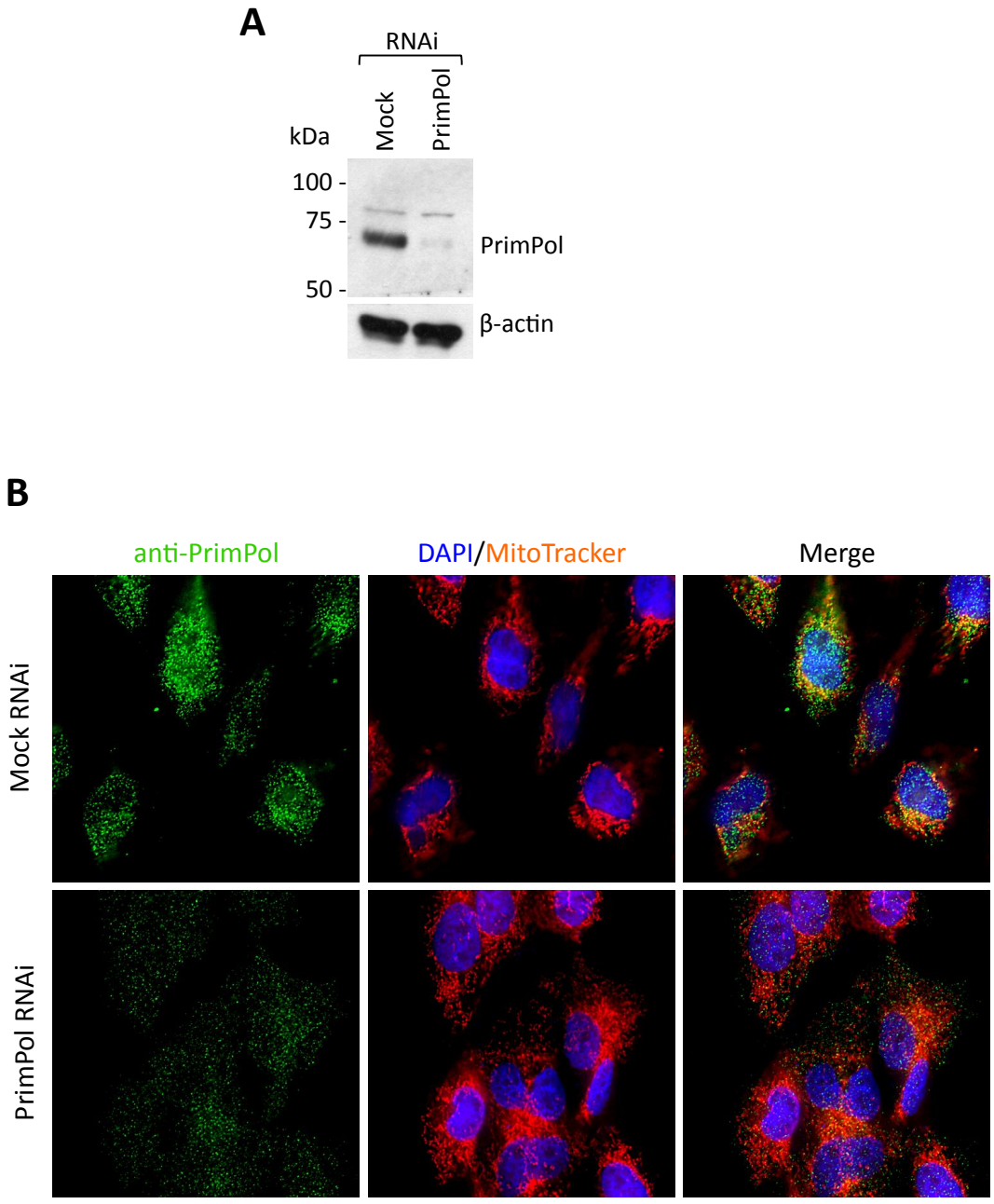


Figure 5.5. Immunostaining with anti-PrimPol is specific to the PrimPol protein.
143B cells were treated with mock or PrimPol RNAi and analysed after 48 hours. (**A**) Cell lysates were prepared and analysed by Western blot with an anti-PrimPol antibody and anti-β-actin as a loading control. (**B**) Cells were fixed and subjected to immunofluorescent analysis with an anti-PrimPol antibody. Prior to fixation cells were incubated with MitoTracker Deep Red to visualise mitochondria (red), and following immunofluorescent staining cells were counterstained with DAPI to visualise nuclei (blue). Images were taken with a 60x objective using identical settings and deconvolved.

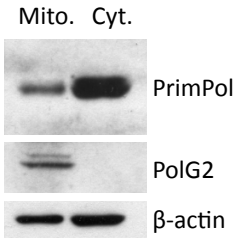


Figure 5.6. PrimPol is present in a crude mitochondrial preparation.
Whole mitochondria were prepared from HEK-293 cells by differential centrifugation with the Mitochondrial Isolation Kit for Cultured Human Cells. The mitochondria were then lysed (mito) and their contents analysed by Western blot alongside the cytoplasmic fraction (cyt), using the indicated antibodies. The mitochondrial DNA polymerase accessory subunit, PolG2, was used as a mitochondrial marker.

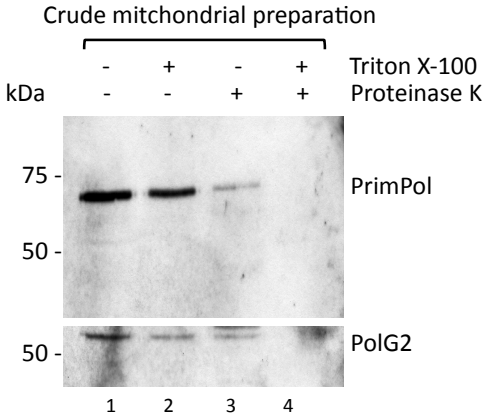


Figure 5.7. PrimPol is retained in protease treated mitochondria.
Whole mitochondria were isolated from U2OS cells and incubated with Proteinase K (0.5 mg/ml) in the absence or presence of detergent Triton X-100 (0.5 %) for 30 minutes on ice. The remaining proteins were then subjected to Western blot analysis with the antibodies indicated. The mitochondrial accessory subunit PolG2 was used as a mitochondrial marker.

mitochondrial membrane and so exposes the proteins inside the mitochondria to Proteinase K, PolG2 was completely degraded (Figure 5.7, lane 4). This indicates that PolG2 was present inside the mitochondria as the double mitochondrial membrane protected PolG2 from proteolysis. A similar result was observed for PrimPol; PrimPol was retained in protease treated mitochondria and completely degraded following addition of Triton X-100 (Figure 5.7). Although, following addition of Proteinase K a reduction in the level of PrimPol protein was observed (Figure 5.7), indicating that some PrimPol protein was present outside the mitochondria. This data is consistent with the immunofluorescent studies confirming that some, but not all, PrimPol is present inside the mitochondria of cultured human cells.

5.4. PrimPol is a nuclear protein, associating with chromatin

Immunofluorescent studies also suggested PrimPol was present in the nuclei of human cells. Nuclear DNA metabolic processes occur in the context of chromatin, and so PrimPol's presence on chromatin was next tested. A simple method to enrich for chromatin is to treat cells with a cytoskeletal buffer containing the detergent Triton X-100, which permeabilises the cell and organelle membranes releasing the soluble proteins inside. A centrifugation step can then separate the Triton X-100 soluble components of the cell from the Triton X-100 insoluble components, the latter containing cellular sub-structures such as chromatin (Kannouche *et al*., 2004). Given PrimPol is a nucleic acid synthetic enzyme the association of PrimPol with chromatin was tested in cells synchronised in S-phase, in addition to an asynchronous cell population. To stall cells in early S-phase aphidicolin was used, a reversible inhibitor of the nuclear replicative (family-B) DNA polymerases (Wist and Prydz, 1979; Wright *et al*., 1994).

SV40-transformed non-diseased human (MRC5) fibroblasts were grown in the presence or absence of aphidicolin for 16 hours, and then the aphidicolin was removed and the cells allowed to grow for a further 3 hours. Analysis of the DNA content of these cells using PI staining and flow cytometery shows that the aphidicolin treatment successfully stalled the cells in early S-phase, and removal of the aphidicolin allowed the cells to progress synchronously through S-phase (Figure 5.8a). These cells were separated into Triton X-100 soluble and insoluble fractions and analysed by Western blot. The chromatin component histone H1 was almost exclusively detected in the insoluble fractions, indicating the enrichment of chromatin in these samples (Figure 5.8b). Very little insoluble PrimPol was observed in asynchronous cells, however, a marked increase of insoluble PrimPol was observed in cells arrested in early S-phase (Figure 5.8b). Release of the early S-phase arrest, allowing cells to progress

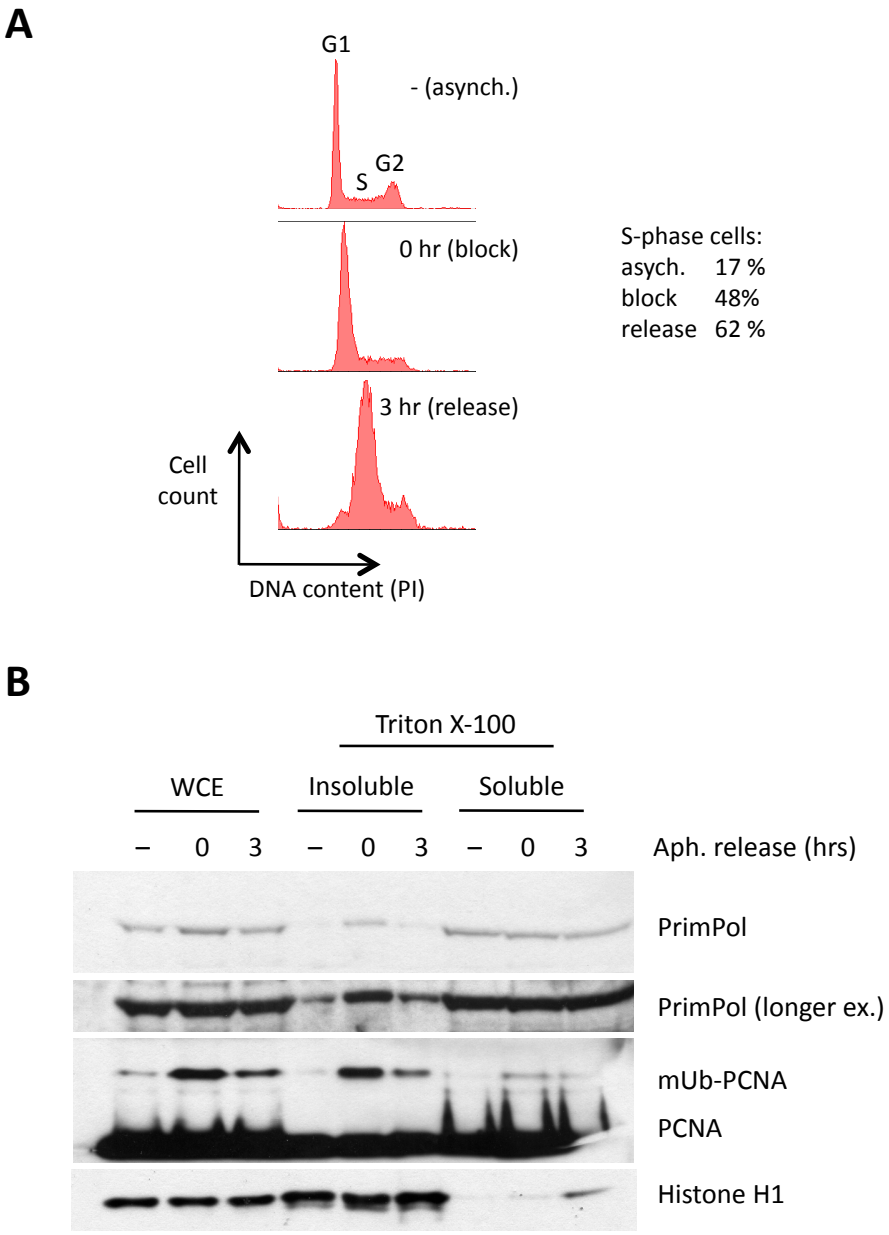


Figure 5.8. PrimPol becomes insoluble following prolonged DNA replication arrest.
SV40-transformed normal (MRC5) fibroblasts were grown in the presence or absence of 2 μg/ml aphidicolin for 16 hours, and then the aphidicolin was removed and cells allowed to grow for a further 3 hours before analysis. (**A**) Cells were fixed, stained with PI, and analysed by flow cytometry to determine the DNA content and the proportion of S-phase cells. (**B**) Cells were resuspended in cytoskeletal buffer containing 0.2 % Triton X-100 for 5 minutes on ice, and then centrifugated to separate the soluble and insoluble fractions. Equal quantities of whole cell extract (WCE), Triton soluble, and Triton insoluble fractions were analysed by Western blot with the antibodies indicated. Detection of Histone H1 serves as a chromatin maker, PCNA and the slower migrating mono-ubiquitinated (mUb)-PCNA serves as a marker of stalled replication forks.

synchronously through S-phase, led to the level of insoluble PrimPol decreasing to almost the level observed in asynchronous cells (Figure 5.8b). This suggests that PrimPol does associate with chromatin in S-phase, but during unperturbed DNA replication the level of chromatin bound PrimPol may be below the detection limit if present at all. The stalling of cells in S-phase with aphidicolin will cause extensive replication fork stalling due to the stalled replicative DNA polymerase functionally uncoupling from the MCM-helicase complex, and under these circumstances chromatin bound PCNA becomes mono-ubiquitinated (Kannouche *et al*., 2004; Chang *et al*., 2006). In line with this an increase in insoluble mono-ubiquitnated PCNA was observed following aphidicolin treatment (Figure 5.8b). As it is under these circumstances when PrimPol's association with chromatin is more apparent, it is possible that PrimPol may play a role at stalled replication forks. Taken together with the immunofluorescent analysis of over-expressed and endogenous PrimPol, these data suggest this novel enzyme is localised to nuclei in addition to the mitochondrion.

5.5. Discussion

DNA is found in two locations within human cells, inside nuclei and inside mitochondria, and it is within these organelles PrimPol was detected. Immunofluorescent detection of endogenous and over-expressed recombinant PrimPol in human cultured cells showed PrimPol was present in the nucleus and cytoplasm, with a portion of the latter being mitochondrial. Sub-cellular fractionation experiments confirmed that PrimPol was present in human mitochondria. PrimPol was also found to associate with chromatin, consistent with its presence in nuclei, and its association became much more apparent when cells were arrested during DNA replication. This dual localisation of PrimPol is achieved despite the apparent lack of localisation target signals, however this is not uncommon, a prime example being the replicative primase Prim1, which “piggy backs” upon its accessory subunit Prim2 in order to localise to the nucleus (Mizuno *et al*., 1996). The implications of the dual localisation of PrimPol to the nucleus and mitochondrion are discussed below.

5.5.1. PrimPol – a novel mitochondrial primase-polymerase in human cells

PrimPol is capable of both primer synthesis and DNA chain elongation *in vitro* (Aidan Doherty and Luis Blanco, unpublished), but determining its *in vivo* activity will no doubt prove difficult. However, regardless of whether PrimPol is a primase, polymerase, or perhaps both *in vivo*, there are a number of important implications of the work presented in this chapter. Firstly, Pol γ is no longer the sole mitochondrial DNA polymerase in human cells, and so the assumption that Pol γ is involved in all DNA synthetic processes within this organelle needs to be re-addressed. It is well established that Pol γ is the mtDNA replicase; inactivation of its 3’-5’ exonuclease proofreading domain in mice results in the accumulation of mutations and deletions in mtDNA (Trifunovic *et al*., 2004), and further, deletion of the catalytic subunit results in early developmental arrest during embryogenesis due to mtDNA depletion, demonstrating that Pol γ is the only DNA polymerase able to maintain mammalian mtDNA (Hance *et al*., 2005). Pol γ also has a fairly well-established role in mitochondrial BER, as this enzyme posses dRP-lyase and gap-filling activity required for short-patch BER (Longley *et al*., 1998) and functionally interacts with helicase/nuclease DNA2 and flap-endonuclease FEN1 in long-patch BER (Liu *et al*., 2008; Zheng *et al*., 2008). The role of Pol γ in other mtDNA metabolic processes however, is less well understood. For instance it is inevitable that some DNA damage will escape repair and encounter the replication machinery, especially as mtDNA replication is constitutive and there is an apparent lack of NER in this organelle (Cline, 2012), which is required for removing bulky replication blocking lesions from the nuclear genome. In the nucleus a group of flexible TLS polymerases exist specifically for this

scenario, and it would be assumed that the high fidelity of Pol γ and proofreading activity would not be compatible with significant TLS activity. Indeed, lesions which create little distortion in the DNA helix such as a templated 8-oxo-G, a common lesion in mitochondria due to the close proximity of ROS production, poses a significant block to human Pol γ (Graziewicz *et al*., 2007). Pol γ itself is also a major target of oxidative damage in mitochondria, and this has been shown to have detrimental effects on its replication and repair capabilities (Graziewicz *et al*., 2002). Bulky helix-distorting DNA lesions such as CPDs caused by UV radiation pose a severe block to Pol γ *in vitro* (Kasiviswanathan *et al*., 2012), and would therefore be expected to stall DNA replication. It is noteworthy that REV3, the catalytic subunit of Pol ζ, and REV1, which are well known players in nuclear TLS, are present within yeast mitochondria (Zhang *et al*., 2006); and it is also noteworthy that PrimPol is absent from all but one fungus. In addition to damage tolerance, the majority of DNA repair mechanisms identified in the nucleus require a DNA polymerase (Lange *et al*., 2011), and it is possible that with regards to mitochondria, this polymerase could by PrimPol. A counter argument for the relative simplicity of mtDNA repair mechanisms, rather than the lack of identified proteins involved, is that mtDNA is polyploid, and so it is not essential to maintain every copy of the mitochondrial genome to allow the mitochondria and cell to function correctly, rather only a small population of mtDNA molecules (reviewed in Larson, 2010).

In addition to being a novel mitochondrial DNA polymerase, PrimPol is also a primase. Primase activity was first detected in human mitochondria almost three decades ago but the enzyme responsible was not identified (Wong and Clayton, 1985). Since then it has become clear that mitochondria do not require a dedicated primase to initiate replication. It has been demonstrated that RNA transcripts synthesised by POLRMT prime leading strand synthesis (Xu and Clayton, 1996), and more recently that POLRMT is also a dedicated lagging strand primase (Wanrooij *et al*., 2008; Fusté *et al*., 2010). Results from these studies, and the fact that the mystery primase identified in 1985 was associated with structural RNA that was required for primase activity (Wong and Clayton, 1986), a feature not yet observed for PrimPol (Luis Blanco and Aidan Doherty, personal communication), strongly suggests that PrimPol is not the long sought after mitochondrial replicative primase. In addition to initiation of DNA synthesis, primases have also been implicated in replication restart. It has been demonstrated in bacteria that when a replication fork encounters a blocking DNA lesion, re-priming can occur downstream of the lesion to facilitate replication progression (Heller and Marians, 2006; Yeeles and Marians, 2012). PrimPol could play a similar role in mitochondria, although this is highly speculative.

5.5.2. PrimPol – a novel primase-polymerase in the nuclei of human cells

In addition to the impact upon our understanding of mtDNA metabolism, PrimPol was also present in the nuclei of human cells. Eukaryotic nuclei already contain a dedicated replicative DNA primase that exists associated to the family-B polymerase Pol α, which is responsible for initiating DNA synthesis at replication origins and Okazaki fragments (reviewed in Muzi-Falconi *et al*., 2003). Given that each subunit of the Pol α-Prim complex is essential for cell viability (Lucchini *et al*., 1987; Foiani *et al*., 1989; Sugino, 1995; Muzi-Falconi *et al*., 2003) it is unlikely that PrimPol plays a similar role in DNA replication. The only other foreseeable role for a primase is to restart DNA synthesis when replication stalls, as demonstrated in *E. coli* (Heller and Marians, 2006; Yeeles and Marians, 2012) and suggested to occur in yeast (Lopez *et al*., 2006; reviewed in Lehmann and Fuchs, 2006). However in *E. coli* it is the replicative DnaG primase that is responsible (Heller and Marians, 2006; Yeeles and Marians, 2012). Additionally, all but one fungus do not contain PrimPol, so again the role would be presumably carried out by Pol α-Prim, as implied by Marini and collegeues (1997).

PrimPol can also be added to the ever-growing list of DNA polymerases present in the nuclei of human cells. As in the nucleus there are a much greater variety of DNA polymerases it is difficult to speculate possible roles, more so than in the relatively simple mitochondrion. Further phenotypic experiments will be required to delineate the role of PrimPol in human cells. However speculated roles in DNA repair and damage tolerance, whether re-priming or TLS, are possible. In bacteria AEP-like enzymes have been shown to be dedicated NHEJ polymerases required for the repair of double-strand breaks in DNA (Della *et al*., 2004; Brissett and Doherty, 2009); this could be a role for PrimPol, although the nuclei of human cells contain an abundance of NHEJ polymerases from the X-family. The observation that PrimPol becomes chromatin bound when DNA replication is stalled, but lesser so when cells are progressing through S-phase, suggests that PrimPol may not be required during unperturbed DNA replication, and may play a role in the cellular response to stalled replication forks. This is investigated further in the following chapter.

Chapter 6

PrimPol, a new player in DNA damage tolerance in human cells

6.1. Introduction

Chapter 5 established that human PrimPol is present in both the nucleus and mitochondrion. The aim of the current chapter was to build upon these findings, together with other data obtained in the Doherty lab, and uncover the specific cellular function(s) of PrimPol. Early experiments in the Doherty lab involved the disruption of the PrimPol gene in chicken DT40 cells, and it became evident that as a result of PrimPol deletion these cells became sensitive to UV-C radiation (Julie Bianchi and Aidan Doherty, unpublished). Exposure of cells to UV-C light results in the formation of covalent linkages between adjacent pyrimidines on the same DNA strand, primarily producing CPDs and 6-4 photoproducts. These DNA lesions significantly distort the DNA helix and therefore pose a significant block to DNA metabolic processes such as DNA replication and transcription (Batista *et al*., 2009; Rastogi *et al*., 2010). DNA polymerases are required for the repair of these lesions during the DNA synthetic step of NER (Lehmann, 2011), and also required for the tolerance of these potentially lethal lesions during DNA replication, which is largely performed by the Y-family DNA polymerases (Sale *et al*., 2012). Primases have also been implicated in DNA damage tolerance, being required to re-prime downstream of DNA lesions thereby allowing continued progression of the replication fork. This has been demonstrated in bacteria (Heller and Marians, 2006; Yeeles and Marians, 2012) and evidence exists that it also occurs in eukaryotes (Lehmann, 1972; Lopez *et al*., 2006; Elvers *et al*., 2011). The following experiments explore the role of human PrimPol in the cellular response to UV-C radiation.

6.2. PrimPol re-localises in the nucleus following UV-C irradiation

I first explored whether UV-C irradiation had an effect on the sub-cellular localisation of PrimPol. This was done first using the Flp-In T-REx-293 cells (HEK-293 derivative) engineered for inducible expression of recombinant PrimPol, which is expressed fused to either a carboxyl-terminal HA or tandem Flag-*Strep*-Tag-II (see section 5.2.1). As discussed in Chapter 5, both the nuclear and mitochondrial localisation is much more apparent when PrimPol is stably over-expressed in these cells (Figure 5.3).

6.2.1. PrimPol assembles into detergent-resistant foci in UV-C irradiated cells

The Flp-In T-REx-293 cells over-expressing recombinant PrimPol were either mock-irradiated or exposed to 30 J/m^2 UV-C light and then analysed by immunofluoresence microscopy following various recovery times (Figure 6.1a). Exposure of cells to UV-C radiation resulted in no obvious change in the observed localisation of PrimPol, whether cells were analysed following a 1, 8, or 24 hour recovery after irradiation (Figure 6.1b). To

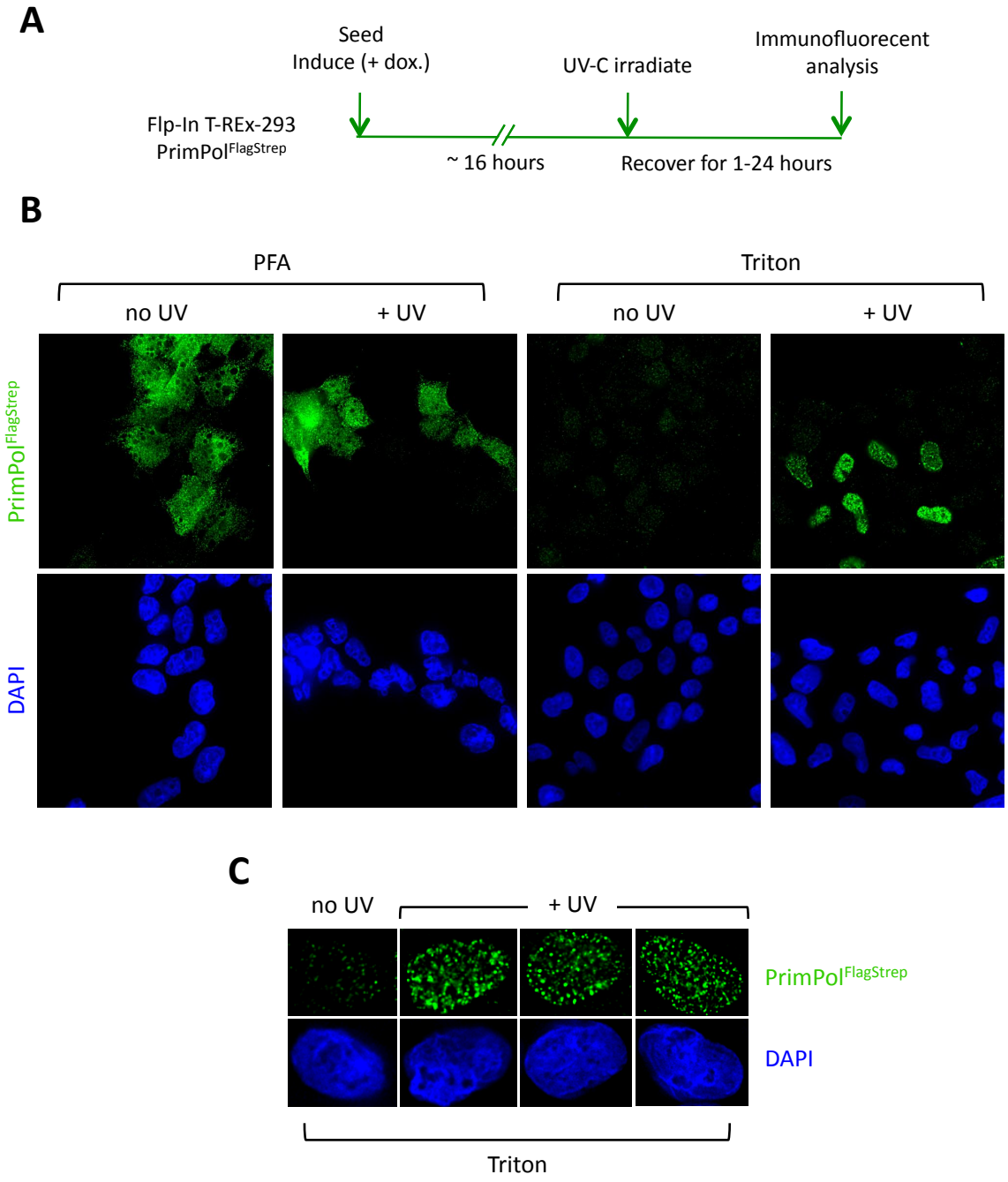


Figure 6.1. PrimPol re-localises into sub-nuclear detergent-resistant foci following UV-C irradiation. (**A**) Schematic detailing experiment. Flp-In T-REx-293 cells engineered for inducible expression of PrimPol^{FlagStrep} were induced by addition of doxycycline (dox) (10 ng/ml) to cell medium for ~16 hours. Cells were either mock or UV-C irradiated (30 J/m^2), allowed to recover for 1-24 hours, and either directly fixed in paraformaldehyde (PFA) or first pre-extracted in a Triton X-100 (0.5 %) containing buffer before immunofluorescent analysis. (**B**) PrimPol was detected using an anti-PrimPol antibody (green) and DNA was counterstained with DAPI (blue) allowing visualisation of nuclei. Representative images following a 1 hour recovery shown, taken with a 60x objective and deconvolved. Similar results were also observed following an 8 or 24 hour recovery. (**C**) Higher magnification images of UV-C induced PrimPol foci in pre-extracted cells. Images taken with a 100x objective and deconvolved.

visualise chromatin associated proteins, as this would be the presumed location of active DNA metabolising enzymes within the nucleus, cells were washed with a Triton X-100 containing buffer prior to fixation (termed pre-extraction). This permeabilises cells *in situ*, removing the majority of soluble proteins leaving only those bound to cellular structures such as chromatin (Kannouche *et al*., 2001). In mock-irradiated cells, over-expressed PrimPol was more or less undetectable following pre-extraction (Figure 6.1b). However, following exposure to 30 J/m^2 UV-C radiation, PrimPol was visible as multiple tiny bright spots within ~40 % of nuclei (Figure 6.1b), numbering between 50 and several hundred per cell (Figure 6.1c). These UV-C induced focal accumulations of PrimPol were observed in 2 independent stable cell lines, being present 1 hour after irradiation and persisting for at least 24 hours. Together, these data show that PrimPol re-localises following exposure of cells to UV-C radiation, and becomes tightly associated with a nuclear sub-structure, which is presumably chromatin.

6.2.2. PrimPol foci assemble in a UV-C dose-dependant manner

To further investigate the relationship between PrimPol foci and UV-C radiation, PrimPol over-expressing cells were exposed to a range of UV-C doses before pre-extraction and immunofluorescent analysis. Exposure of cells to 5 J/m^2 UV-C radiation resulted in very little to no detectable PrimPol foci (Figure 6.2a). Following 10 J/m^2 UV-C radiation around 10 % of cells presented focal PrimPol, which increased to almost 40 % following irradiation with 30 J/m^2 UV-C (Figure 6.2a). This appeared to be the upper limit of PrimPol focal cells. These data indicate that PrimPol accumulates into sub-nuclear detergent-resistant foci in a UV-C dose dependant manner.

6.2.3. PrimPol foci do not assemble following exposure to ionising radiation

To test whether PrimPol's re-localisation following UV-C irradiation was a non-specific cellular response to DNA damage, the distribution of over-expressed PrimPol was monitored following exposure of cells to ionising radiation, which produces both single-strand and double-strand breaks in DNA (Ward, 1975). Immunofluorescent analysis of pre-extracted cells revealed no focal accumulation of PrimPol following exposure to 5 Gy X-radiation, with little to no PrimPol being detected following recovery times varying from 1-24 hours (Figure 6.3). This suggests that focal accumulation of PrimPol only occurs following the production of specific DNA lesions, such as those caused by UV-C irradiation.

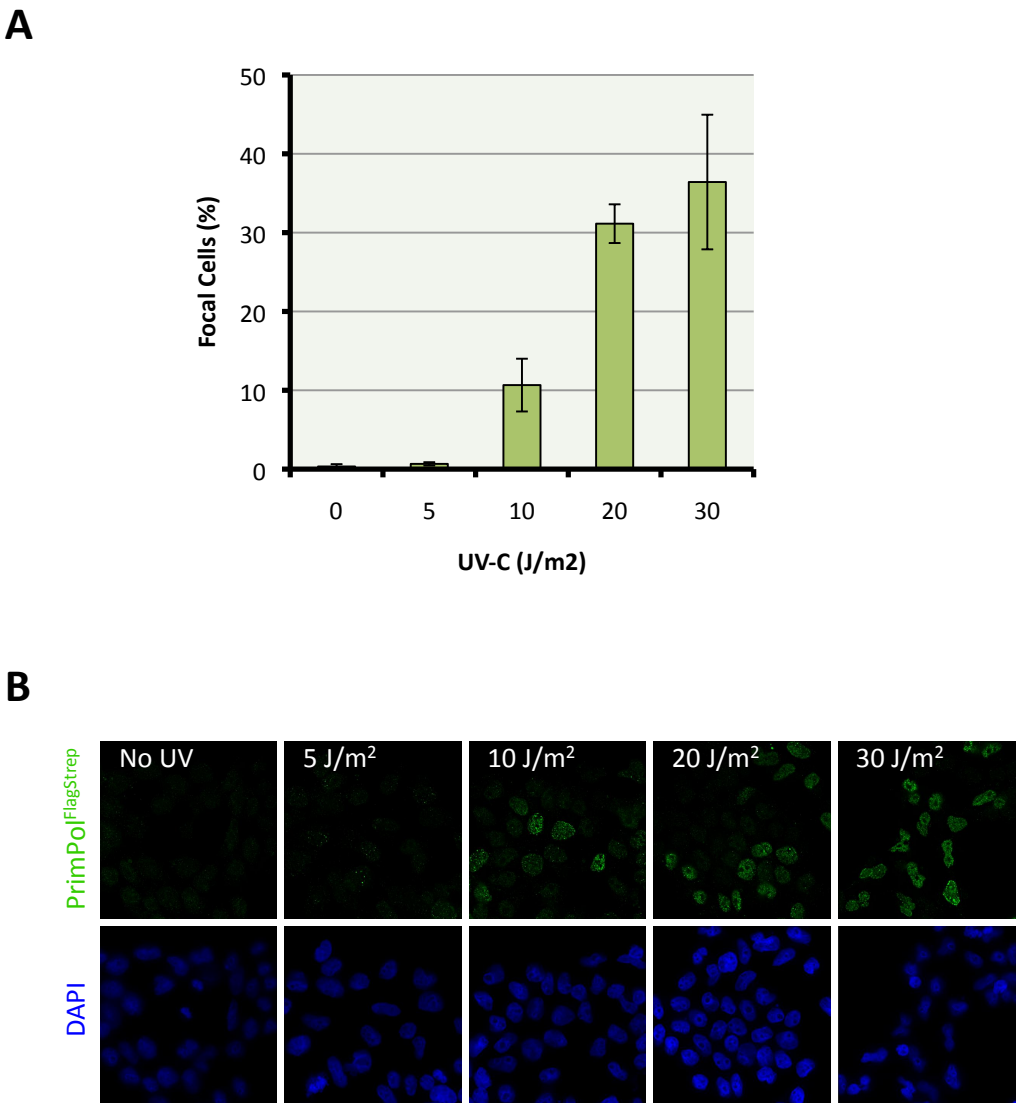


Figure 6.2. The proportion of PrimPol focal cells was dependant upon the dose of UV-C irradiation. (**A**) Flp-In T-REx-293 cells engineered for inducible expression of PrimPol^{FlagStrep} were induced by addition of doxycycline (10 ng/ml) to cell medium for ~16 hours, irradiated with UV-C doses indicated, and allowed to recover for 8 hours prior to pre-extraction and immunofluorescene staining with anti-PrimPol and counterstaining with DAPI. The proportion of cells in which PrimPol was localised into foci was determined. At least 200 cells were counted for each dose and the experiment repeated 3 times, error bars indicate one standard deviation. (**B**) Panel of representative immunofluorescent images of PrimPol focal cells at various UV-C doses, anti-PrimPol (green) and DAPI (blue). Images taken with a 60x objective and deconvolved.

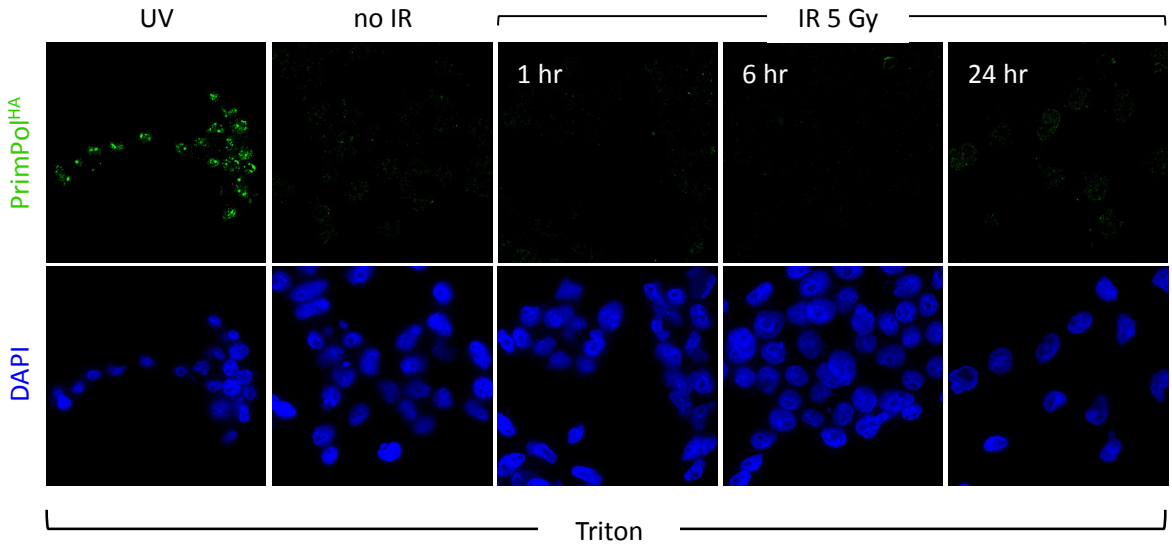


Figure 6.3. PrimPol did not assemble into foci following exposure to ionising radiation.
Flp-In T-REx-293 cells engineered for inducible expression of PrimPol^{HA} were induced by addition of doxycycline (10 ng/ml, 16 hours) and either mock or X-ray (5 Gy) irradiated and allowed to recover for the times indicated before pre-extraction and immunofluorescent staining with anti-HA (green) and counterstaining with DAPI (blue). As a positive control cells were also irradiated with UV-C (20 J/m^2). Images taken with a 60x objective and deconvolved.

6.3. Does PrimPol re-localise to stalled replication forks?

Collision of a replicative DNA polymerase with a DNA lesion can cause functional uncoupling of the MCM helicase complex and DNA polymerase, which in turn produces long tracts of unwound single-stranded DNA that becomes coated with the single-stranded DNA binding protein RPA (Byun *et al*, 2005). These long stretches of single-stranded DNA have actually been visualised *in vivo* by electron microscopy in yeast (Lopez *et al*., 2006). Following treatment of cells with replication fork blocking agents RPA re-localises into detergent-resistant sub-nuclear foci (Vassin *et al*, 2004; Robison *et al*, 2004) thought to represent these long stretches of RPA-coated single-stranded DNA rather than sites of DNA repair such as NER (Despras *et al*., 2010; Pathania *et al*, 2011; Diamant *et al*., 2012). It was therefore investigated whether PrimPol co-localised with RPA following UV-C irradiation, as this would suggest PrimPol was present at replication forks stalled at UV photoproducts. Following exposure of cells over-expressing PrimPol to UV-C radiation, all cells that were focal for RPA2 were also focal for PrimPol (Figure 6.4a), suggesting PrimPol foci occur during S-phase. Furthermore, a proportion of PrimPol foci co-localised with RPA foci (Figure 6.4a and b), suggesting PrimPol was localised at a sub-population of stalled replication forks.

To investigate this further, cells over-expressing PrimPol were UV-C irradiated and during a 1 hour recovery, pulse labelled with EdU (5-ethynyl-2'-deoxyuridine) (Figure 6.5a), a nucleoside analogue of thymidine that would be incorporated into DNA during DNA synthesis. Therefore immunofluorescent detection of EdU would allow determination of S-phase cells in an asynchronous population, and it was tested whether these cells contained focal PrimPol following UV-C irradiation. The vast majority of cells focal for PrimPol also contained EdU staining, and were therefore S-phase cells, following exposure to 15 J/m^2 UV-C radiation (Figure 6.5b). This suggests that UV-C induced focal accumulation of PrimPol occurs in S-phase cells. Although this last result is preliminary, taken together with the co-incidence of PrimPol and RPA foci, these data suggest that re-localisation of PrimPol in UV-C irradiated cells occurs during S-phase, and that a proportion of PrimPol may localise to replication forks stalled at UV photoproducts. This would be consistent with the previous observation that the upper limit of PrimPol focal cells in an asynchronous population was ~40 % (Figure 6.2a), which would correspond to the proportion of S-phase cells in an asynchronous population.

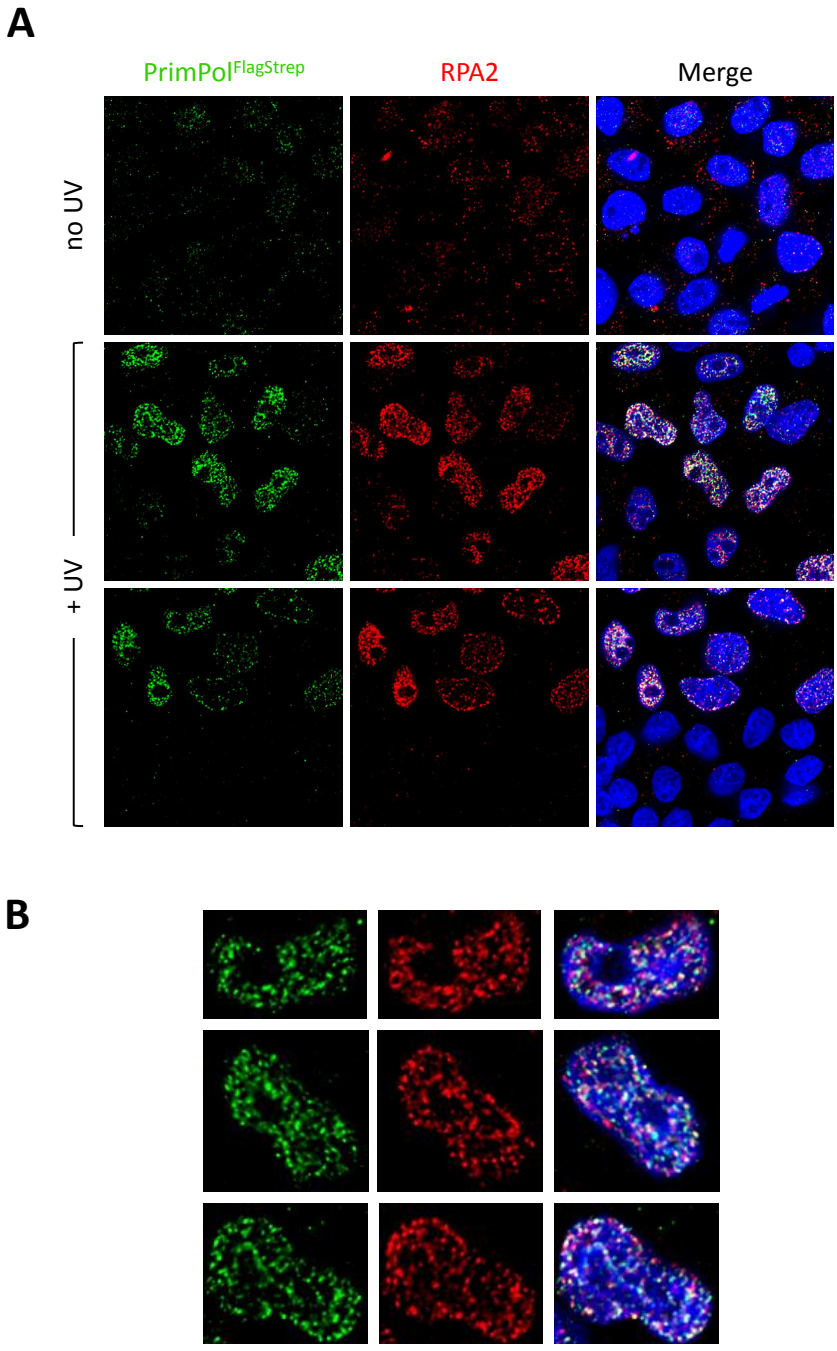


Figure 6.4. PrimPol focal cells following UV-C irradiation are also focal for RPA, and a portion of these foci co-localise. (**A**) Flp-In T-REx-293 cells engineered for inducible expression of PrimPol^{FlagStrep} were induced by addition of doxycycline (10 ng/ml, 16 hours), mock or UV-C irradiated (30 J/m^2), and allowed to recover for 1 hour before pre-extraction (0.5 % Triton X-100) and immunofluoresent staining with anti-PrimPol (green), anti-RPA2 (red), and counterstaining with DAPI (blue). Images taken with a 100x objective and deconvovled (**B**) Panel of higher magnification images showing focal PrimPol and RPA in the nuclei of UV-C irradiated cells. Images taken with a 100x objective, deconvolved, and digitally magnified using OMERO software.

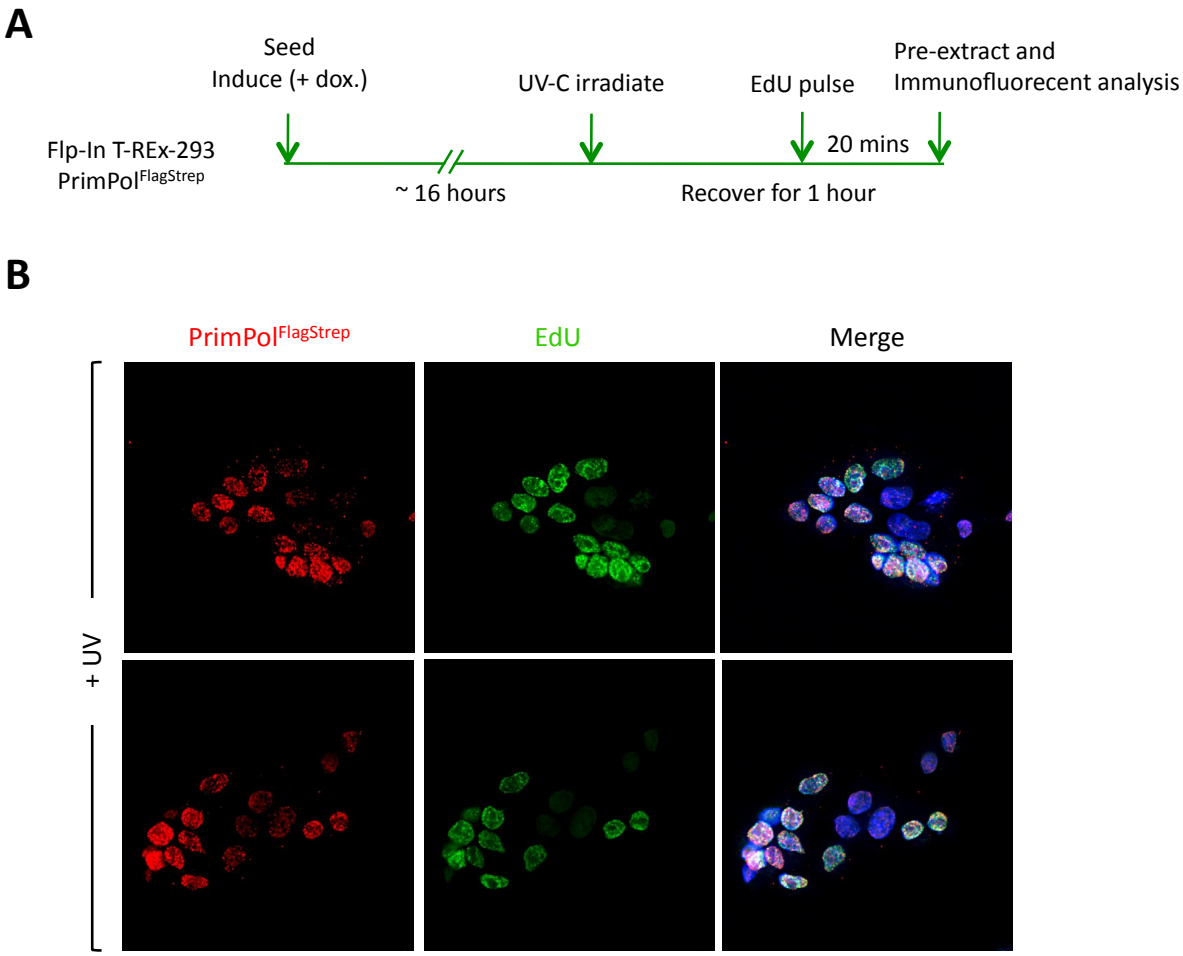


Figure 6.5. UV-C induced PrimPol foci occur in S-phase cells (preliminary)**.**
(**A**) Schematic detailing experiment. Flp-In T-REx-293 cells engineered for inducible expression of PrimPol^{FlagStrep} were induced by addition of doxycycline (dox) (10 ng/ml, 16 hours). Cells were UV-C irradiated (15 J/m^2) and after a 40 minute recovery, pulsed with EdU (10 mM) for 20 minutes before pre-extraction and fixation. (**B**) Immunofluorescent staining with anti-PrimPol (red), anti-EdU (green) and counterstaining with DAPI (blue). Images taken with a 60x objective and deconvolved. Experiment only performed once.

6.4. PrimPol associates tightly with chromatin in UV-C irradiated cells

The accumulation of PrimPol into detergent resistant sub-nuclear foci would be consistent with PrimPol being associated to chromatin. To test this, chromatin was prepared from UV-C irradiated cells using the method described in the previous chapter (section 5.4). Briefly, cells were resuspended in a cytoskeletal buffer containing Triton X-100 and the resulting soluble and insoluble fractions were then separated by centrifugation. Cellular sub-structures such as chromatin are enriched in the Triton insoluble fraction.

6.4.1. Recombinant PrimPol associates with chromatin in UV-C irradiated cells

The Flp-In T-REx-293 cells over-expressing recombinant PrimPol were either mock-irradiated or exposed to 30 J/m^2 UV-C radiation, and following a 1 or 8 hour recovery, the Triton soluble and insoluble fractions were prepared and analysed by Western blot (Figure 6.6a). The chromatin component Histone H1 was exclusively detected in the insoluble fraction (Figure 6.6b), indicating this fraction was enriched for chromatin. Mono-ubiquitinated PCNA, visible following UV-C irradiation, was also predominantly in the insoluble fraction reflecting the role of this protein at replication forks stalled at UV photoproducts (Kannouche *et al*., 2004; Chang *et al*., 2006). In mock-irradiated cells PrimPol was detected in the insoluble fraction, albeit at low levels (Figure 6.6b), however following exposure of cells to 30 J/m^2 UV-C radiation, the level of insoluble PrimPol increased substantially (Figure 6.6b). Given that the total level of PrimPol did not change following irradiation, whilst the level of soluble PrimPol decreased (Figure 6.6b), this is consistent with a re-distribution of the PrimPol cellular pool following irradiation, which may not be surprising given PrimPol was ectopically expressed. The mtDNA replicase Pol γ was also present in the insoluble fraction, and actually became more insoluble following irradiation (Figure 6.6b). However, given that PrimPol assembles into sub-nuclear foci, the increase of insoluble PrimPol following UV-C irradiation most likely reflects association of PrimPol with a nuclear sub-structure, presumably chromatin.

To determine unequivocally PrimPol's association with chromatin, the Triton X-100 insoluble fraction from UV-C irradiated cells was treated with DNase, and further separated into soluble and insoluble fractions by centrifugation (Figure 6.7a). Consistent with a previous report, this completely solubilised mono-ubiquitinated PCNA (Figure 6.7b), which will be bound to DNA at stalled replication forks (Kannouche *et al*, 2004). Similarly, PrimPol was almost completely solubilised by DNase treatment (Figure 6.7b), confirming that PrimPol becomes insoluble following UV-C irradiation due to its association with DNA.

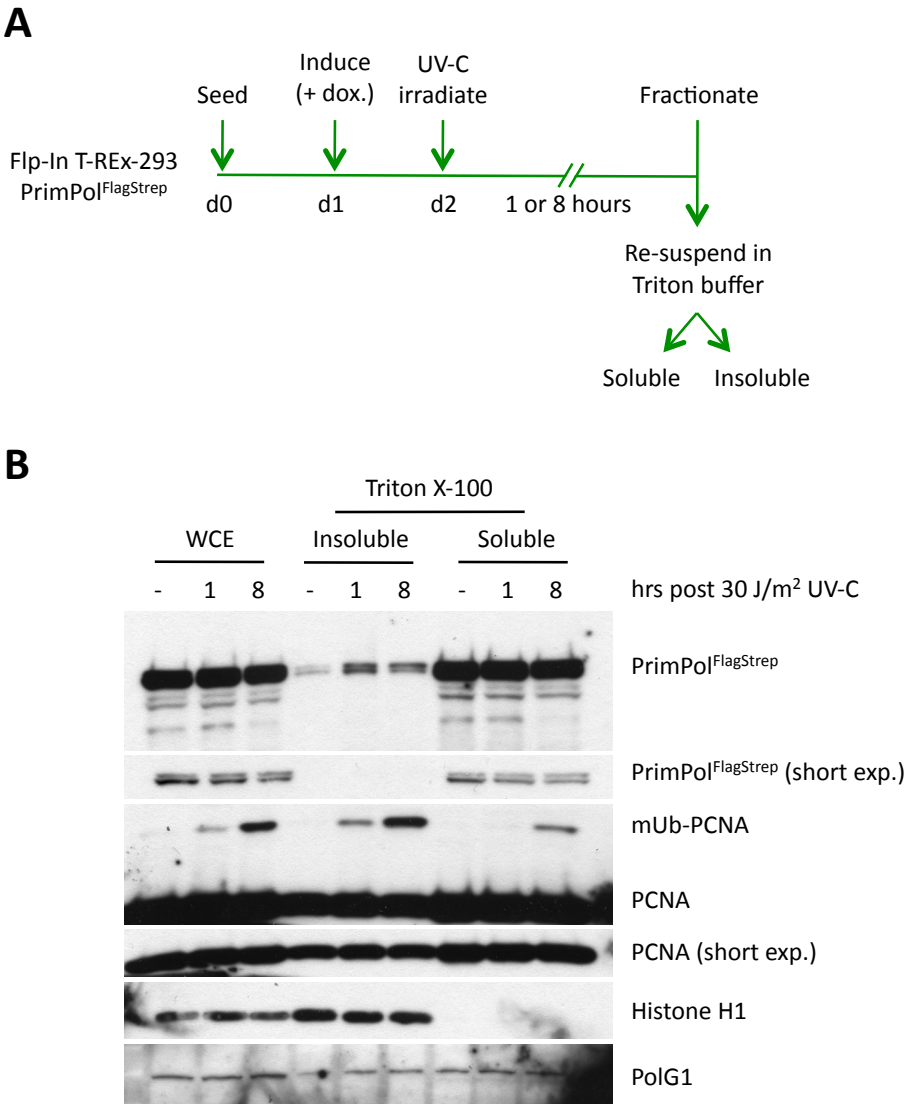


Figure 6.6. Over-expressed PrimPol becomes Triton insoluble following UV-C irradiation.
(**A**) Schematic detailing experiment. Flp-In T-REx-293 cells engineered for inducible expression of PrimPol^{FlagStrep} were induced by addition of doxycycline (dox) (10 ng/ml, 16 hours) and either mock (-) or UV-C (30 J/m^2) irradiated and allowed to recover for 1 or 8 hours before separation into Triton soluble and insoluble fractions. (**B**) Equivalent amounts of whole cell extract (WCE) and the triton soluble and insoluble fractions were resolved by SDS-PAGE and analysed by Western blot with the antibodies indicated. PrimPol^{FlagStrep} was detected with an anti-PrimPol antibody, detection of Histone H1 serves as a maker of chromatin, PCNA and the slower migrating mono-ubiquitinated (mUb)-PCNA serves as a marker UV-C induced DNA damage, and PolG1 as a mitochondrial marker. Representative blot of at least 3 experiments shown.

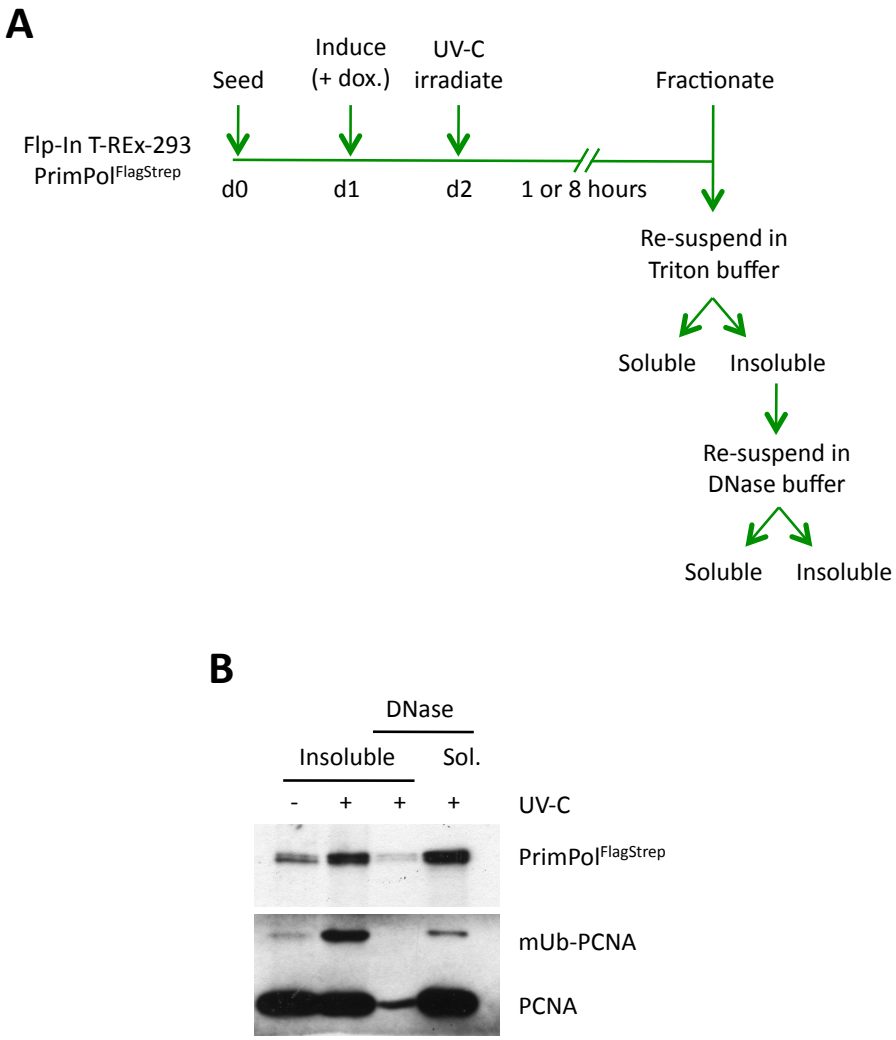


Figure 6.7. Over-expressed PrimPol associates with chromatin following UV-C irradiation. (**A**) Schematic detailing experiment. Flp-In T-REx-293 cells engineered for inducible expression of PrimPol^{FlagStrep} were induced by addition of doxycycline (dox) (10 ng/ml, 16 hours) and either mock (-) or UV-C (30 J/m^2) irradiated and allowed to recover for 1 or 8 hours before separation into Triton soluble and insoluble fractions. The Triton insoluble fraction was resuspended in a DNase (Benzonase) containing buffer to separate into a DNase soluble and insoluble fraction. (**B**) Equivalent amounts of the Triton insoluble and the DNase treated fractions were analysed by Western blot. PrimPol^{FlagStrep} was detected with an anti-PrimPol antibody, detection of PCNA and the slower migrating mono-ubiquitinated (mUb)-PCNA serves as a marker UV-C induced DNA damage, and a protein known to be associated with chromatin. Representative blot shown of at least 3 experiments shown.

6.4.2. Endogenous PrimPol becomes Triton insoluble in UV-C irradiated cells

The previous immunofluorescent and sub-cellular fractionation experiments were performed using cells over-expressing recombinant PrimPol, as this increased the likelihood of detecting changes in PrimPol localisation. It is important to determine whether endogenous PrimPol also re-localises following UV-C irradiation. To that end, SV40-transformed non-diseased (MRC5) fibroblasts were separated into Triton X-100 soluble and insoluble fractions following UV-C irradiation and analysed by Western blot. Both Histone H1 and mono-ubiquitinated PCNA were enriched in the insoluble fraction (Figure 6.8). Very little PrimPol was insoluble in mock-irradiated cells, however following UV-C irradiation, a substantial increase of insoluble PrimPol was visible (Figure 6.8), consistent with the observations of over-expressed PrimPol. Also, the total level of PrimPol did not change following irradiation (Figure 6.8) suggesting there is a re-distribution of the PrimPol cellular pool. Taken together, these data suggest that following exposure of cells to UV-C irradiation the soluble pool of PrimPol re-localises and tightly associates with chromatin at a number of discrete locations, possibly replication forks stalled at UV photoproducts, where presumably it is required for a nucleic acid synthetic process.

6.5. RNAi depletion of PrimPol in normal and XP-V patient cells

Given the re-localisation of PrimPol onto chromatin following UV-C irradiation, it was speculated that PrimPol could play a role in the tolerance of UV photoproducts. To explore this possibility, experiments were performed on SV40-transformed normal (MRC5) and XP-V patient derived (XP30RO) fibroblasts treated with PrimPol RNAi. XP-V is an inherited disorder characterised by sunlight hypersensitivity and a high incidence of skin cancers due to a mutation in the gene encoding Pol η (Masutani *et al*., 1999b; Johnson *et al*., 1999b), the Y-family polymerase which is highly specialised to accurately and efficiently bypass the most common UV photoproduct (Johnson *et al*., 1999a; McCulloch *et al*., 2004; Silverstein *et al*., 2010; Biertümpfel *et al*., 2010). By using cells that are deficient in a major UV DNA damage tolerance pathway, the role of PrimPol within the cell may become more apparent. PrimPol was efficiently depleted using RNAi in both normal and XP-V fibroblasts (Figure 6.9a), and it is noteworthy that a significant growth defect was observed in both cell lines, and this growth defect was much more pronounced in XP-V cells (Figure 6.9b).

6.5.1. PrimPol is required for the tolerance of UV-C induced DNA damage

Defects in the tolerance of UV-induced DNA damage lead to enhanced activation of the intra-S checkpoint following exposure of cells to UV-C radiation, which was shown not to

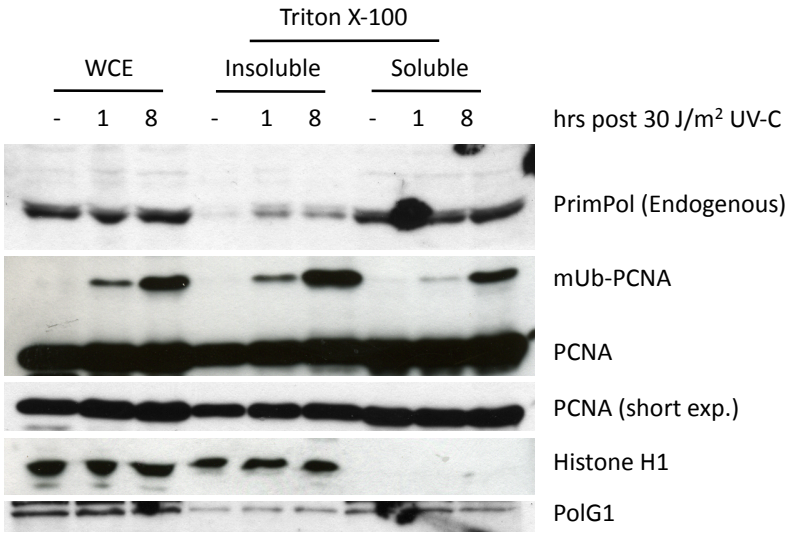


Figure 6.8. Endogenous PrimPol becomes Triton-insoluble following UV-C irradiation.
SV40-transformed non-diseased MRC5 fibroblasts were either mock (-) or UV-C irradiated (30 J/m^2) and following recovery for the indicated time, were separated into Triton soluble and insoluble fractions. Equivalent amounts of whole cell extracts (WCE), Triton soluble, and insoluble fractions were resolved by SDS-PAGE and analysed by Western blot with the antibodies indicated. Histone H1 serves as a chromatin maker, PCNA and the slower migrating mono-ubiquitinated (mUb)-PCNA serves as a marker UV-C induced DNA damage, and PolG1 as mitochondrial marker.

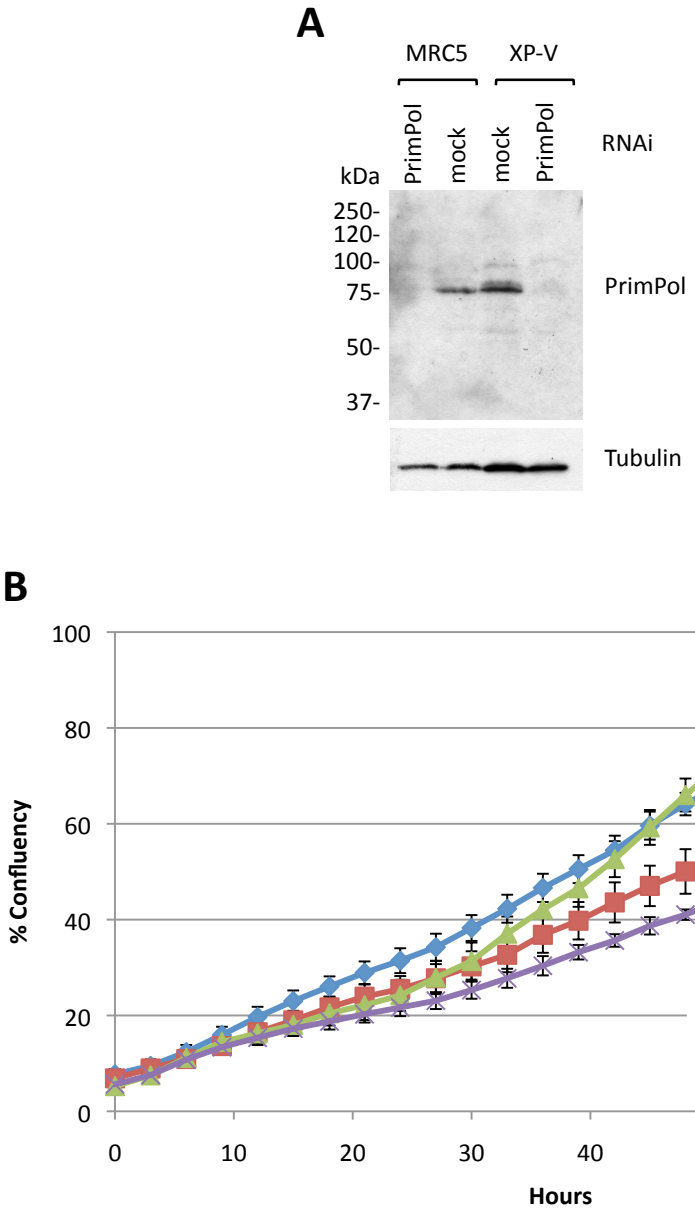


Figure 6.9. PrimPol depletion causes a proliferative defect, which is more pronounced in XP-V cells. (**A**) SV40-transformed normal (MRC5) and XP-V (XP30RO) fibroblasts were mock or PrimPol RNAi treated when seeded, and following 48 hours, cell lysates prepared and analysed by Western blot with the antibodies indicated. (**B**) Immediately after RNAi transfection and seeding, cells were placed in an Incucyte phase contrast microscope and cell density monitored at 3 hourly intervals over 72 hours. Data are representative of 3 independent experiments in which 9 regions were analysed per sample, error bars denote standard error.

be the case for defects in the NER pathway required to remove UV photoproducts (Bomgarden *et al.*, 2006). RPA-coated single-stranded DNA is an initial trigger of the intra-S checkpoint, produced when the replisome encounters a UV photoproduct and functional uncoupling of leading and lagging strand synthesis occurs (Chang and Cimprich, 2009). Normal and XP-V fibroblasts were treated with two rounds of mock or PrimPol RNAi and then exposed to 10 J/m^2 UV-C radiation. Following a 6 hour recovery these cells were detergent extracted and subjected to immunofluorescent analysis to detect RPA2 (Figure 6.10a). In mock-irradiated cells very little to no RPA was detected (Figure 6.10b), consistent with no great perturbation of DNA replication. UV-C irradiation of normal fibroblasts resulted in ~20 % of cells containing sub-nuclear detergent-resistant RPA foci, which increased significantly to ~30 % in PrimPol RNAi treated cells (Figure 6.10b). This indicates a greater proportion of chromatin-associated RPA in UV-C irradiated PrimPol depleted cells, which would be consistent with increased replication fork stalling at UV photoproducts and therefore more RPA-coated single-stranded DNA. DNA replication in UV-C irradiated XP-V cells was already substantially perturbed, as to be expected (Despras *et al*., 2010; Elvers *et al*., 2012), and so no significant increase from ~30 % of focal RPA cells was observed following PrimPol RNAi (Figure 6.10b). This is because detergent-resistant RPA foci following UV-C irradiation occur in S-phase cells, and in an asynchrounous population, 30-40 % of cells will be in S-phase.

The formation of RPA-coated single-stranded DNA recruits and activates the ATR kinase that phosphorylates its main effector kinase Chk1 (Kaufmann, 2010). Previous studies of XP-V cells found that a low dose of 2 J/m^2 UV-C led to an S-phase delay and over-activation of the intra-S checkpoint (Bullock *et al*., 2001; Cordeiro-Stone *et al*., 2002; Despras *et al*, 2010). For this reason, both normal and XP-V fibroblasts treated with mock or PrimPol RNAi were exposed to 2 J/m^2 UV-C radiation and the level of Chk1 phosphorylation on serine 345 determined, as this reflects the level of Chk1 activation (Capasso *et al*., 2002). RNAi depletion of PrimPol resulted in over-activation of Chk1 in UV-C irradiated normal fibroblasts, which was visible 3 hours after irradiation and had returned almost to basal levels following 24 hours (Figure 6.11b). XP-V cells showed substantially more Chk1 activation following UV-C irradiation (Figure 6.11b), indicating a greater perturbation of DNA replication in the absence of Pol η than in the absence of PrimPol. PrimPol depletion in XP-V cells did not increase the initial level of Chk1 activation, however Chk1 activation was persistent in these cells, remaining at a similar level 24 hours after irradiation (Figure 6.11b). Consistent with this, 24 hours after UV-C irradiation an increase in hyperphosphorylated RPA2 was also observed, which is a direct downstream target of ATR, and is phosphorylated following perturbation of DNA

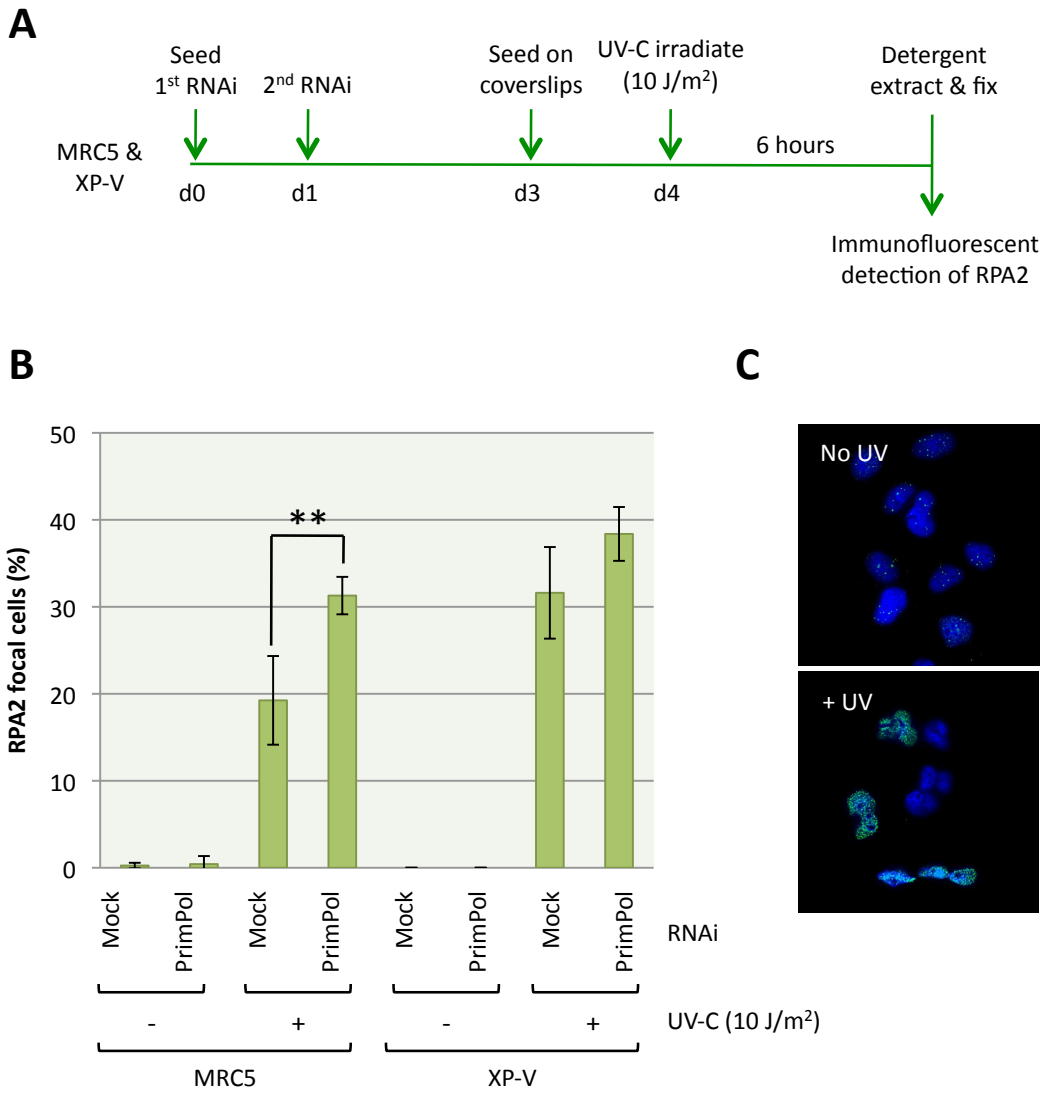


Figure 6.10. Depletion of PrimPol results in increased chromatin-bound RPA in UV-C irradiated normal, but not XP-V, patient cells. (**A**) Schematic detailing experiment. SV40-transformed normal (MRC5) and XP-V (XP30RO) fibroblasts were treated with two consecutive rounds of either mock or PrimPol RNAi and at 72 hours seeded on glass coverslips. The next day (d) cells were either mock or UV-C (10 J/m^2) irradiated and following a 6 hour recovery, Triton X-100 (0.5 %) extracted and subjected to immunofluorescent analysis with anti-RPA2 and DAPI counterstaining. (**B**) The proportion of cells with focal RPA2 (> 20 foci per nuclei) were determined, at least 200 cells were counted from each sample, error bars denote standard deviation of 3 experiments. Significance determined with the two-tailed T-test (** < 0.01). (**C**) Representative image of RPA (green) focal cells seen following UV-C irradiation. Images taken with a 60x objective and deconvolved.

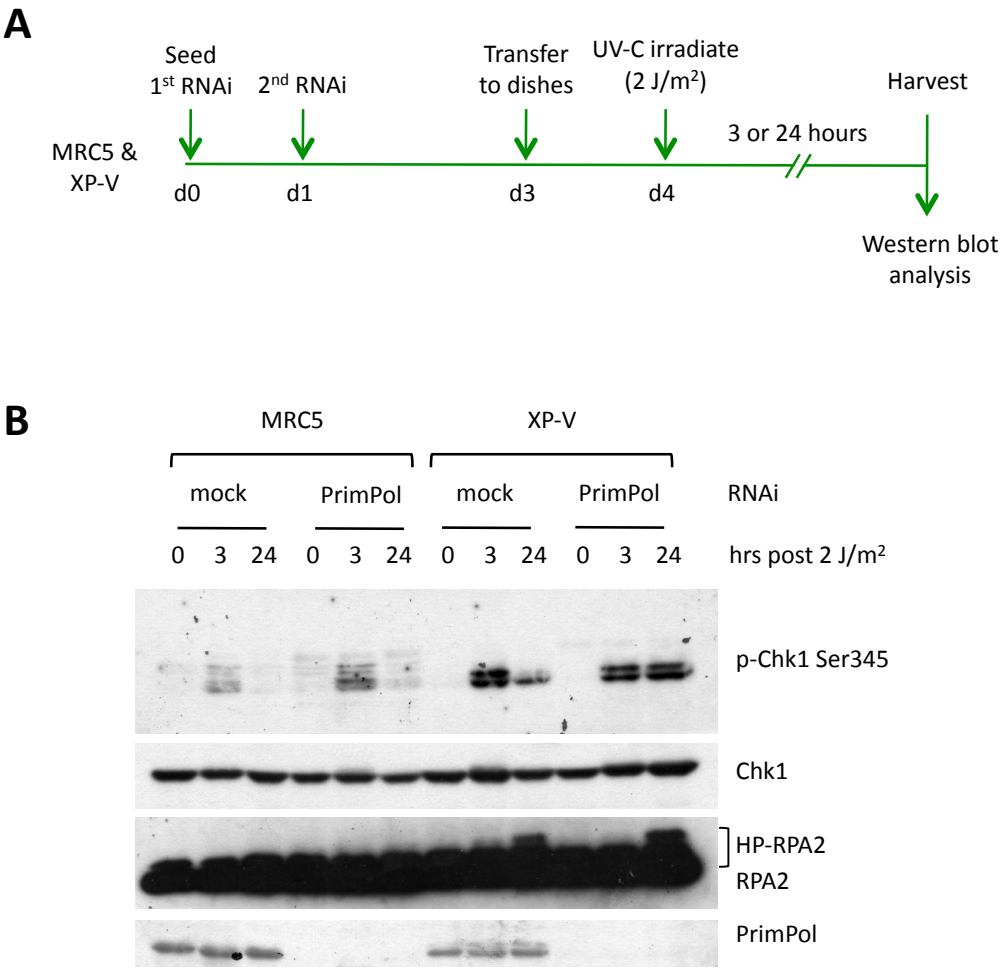


Figure 6.11. PrimPol depletion results in over-activation of the intra-S-checkpoint in UV-C irradiated cells. (**A**) Schematic detailing experiment. SV40-transformed normal (MRC5) and XP-V (XP30RO) fibroblasts were treated with two consecutive rounds of either mock or PrimPol RNAi and at 72 hours transferred to dishes. The next day (d) cells were either mock or UV-C (2 J/m^2) irradiated and following a 3 or 24 hour recovery, whole cell lysates prepared. (**B**) Western blot analysis of whole cell lysates with the antibodies indicated. Hyper-phosphorylation of RPA2 (HP-RPA2) is indicated.

replication (Olson *et al*., 2006) (Figure 6.11b). Taken together, these data indicate that PrimPol is required for the proper cellular response to low dose UV-C radiation in normal and XP-V cells. These data suggest that in the absence of PrimPol, UV-C irradiation produces a greater number of stalled replication forks that in turn over-activate the intra-S checkpoint. Further, these data implicate PrimPol in the tolerance of UV-induced DNA damage.

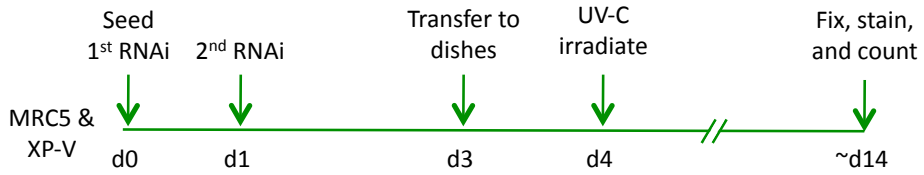
6.5.2. PrimPol protects XP-V, but not normal cells, from UV-C cytotoxicity

It was next investigated whether PrimPol contributed to cell survival following UV-C irradiation in both normal and XP-V fibroblasts, using the clonogenic survival assay. This tests the ability of a cell to survive a particular cytotoxic threat and continue to proliferate indefinitely, forming a large colony of cells visible by the naked eye. The following survivals were performed in the absence of a low concentration of caffeine, which is widely used in the literature and renders XP-V cells hypersensitive to UV-C killing (Arlett *et al*., 1975). RNAi depletion of PrimPol did not sensitise cells to UV-C killing (Figure 6.12), and XP-V cells treated with mock RNAi were only mildly sensitive (Figure 6.12), consistent with previous reports (Arlett *et al*., 1975). However, PrimPol depletion in XP-V cells rendered them synergistically sensitive to UV-C irradiation, decreasing the surviving fraction up to 4-fold (Figure 6.12). This demonstrates that PrimPol protects XP-V cells from UV cytotoxicity, and operates in a damage tolerance pathway that is non-epistatic with Pol η.

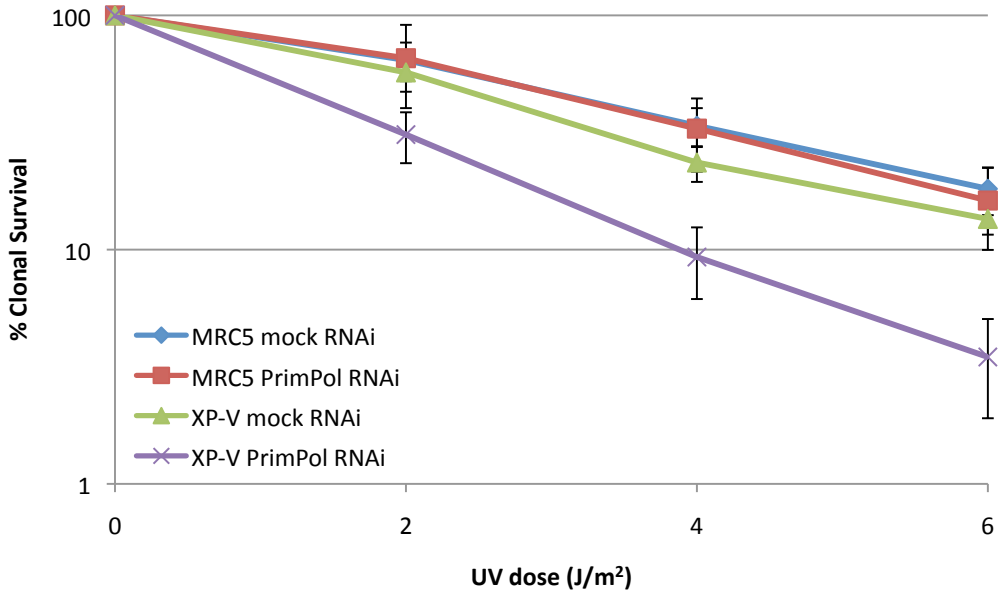
6.6. PrimPol associates with chromatin at discrete foci following nucleotide deprivation

UV-C irradiation of cells is a potent inducer of replication fork stalling due to the physical blockage the UV photoproducts present to the replicative DNA polymerases. Replication forks can also stall in the absence of physical blockage, by inhibition of DNA replication enzymes. Hydroxyurea (HU) inhibits ribonucleotide reductase (RNR) and as a result lowers the cellular pool of dNTPs, thus stalling DNA replication (Young and Hodas, 1964; Moore, 1969). It was next investigated whether treatment of cells with HU resulted in the re-distribution of PrimPol, as previously observed for UV-C irradiation. Flp-In T-REx-293 cells over-expressing recombinant PrimPol were treated with 10 mM HU for 6 hours, before detergent extraction and immunofluorescent analysis. As observed with UV-C irradiation, PrimPol re-localised to sub-nuclear detergent-resistant foci after HU treatment (Figure 6.13a). These foci were also co-incident with focal RPA (Figure 6.13a). Separation of HU treated cells into Triton X-100 soluble and insoluble fractions revealed

A



B



C

	MRC5		XP-V	
J/m²	mock RNAi	PrimPol RNAi	mock RNAi	PrimPol RNAi
0	100 (± 0)	100 (± 0)	100 (± 0)	100 (± 0)
2	64.5 (± 12.1)	65.6 (± 25.5)	57.1 (± 9.9)	31 (± 7.6)
4	33.8 (± 6.4)	32.8 (± 11.3)	23.5 (± 4.1)	9.3 (± 3.2)
6	18.2 (± 4.2)	16.2 (± 6.2)	13.5 (± 2)	3.5 (± 1.6)

Figure 6.12. PrimPol protects XP-V cells from UV cytotoxicity.
(**A**) Schematic detailing experiment. SV40-transformed normal (MRC5) and XP-V (XP30RO) fibroblasts were treated with two consecutive rounds of either mock or PrimPol RNAi and at 72 hours transferred to 10 cm dishes. The following day (d) the cells were mock or UV-C irradiated with the doses indicated and allowed to recover for ~10 days. Colonies were then fixed, stained, and counted. (**B**) UV-C clonogenic survival. The average percentage survival at each UV-C dose is plotted, error bars denote standard deviation of 3 experiments. (**C**) Percentage survival values for graph in (B), number in brackets indicates standard deviation.

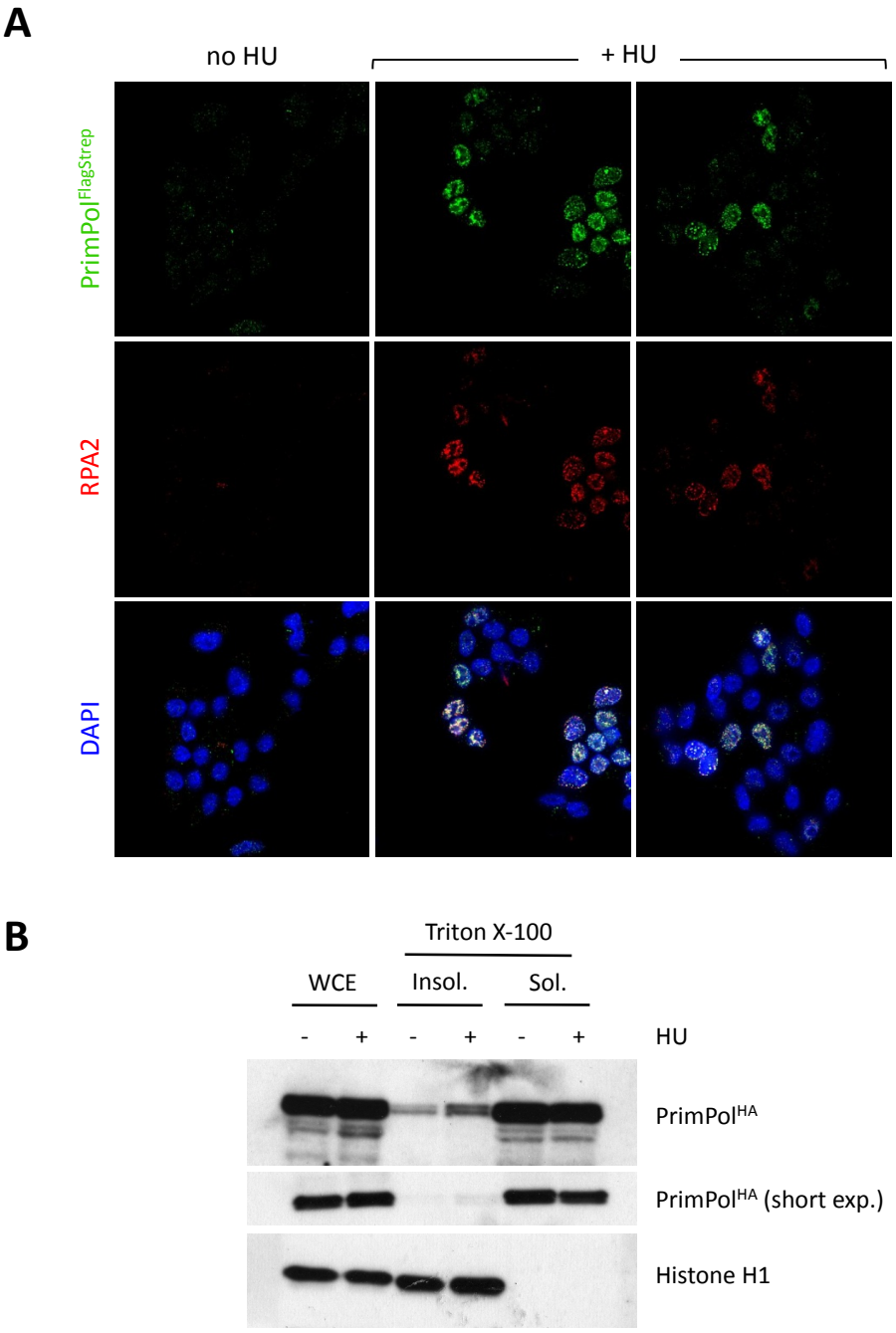


Figure 6.13. PrimPol re-localises following HU treatment, assembling into foci and associating with chromatin. Flp-In T-REx-293 cells engineered for inducible expression of PrimPol^{FlagStrep} were induced by addition of doxycycline (16 hours, 10 ng/ml) and then supplemented with HU (10 mM) for 6 hours before analysis. (**A**) Cells were pre-extracted (Triton X-100 0.5 %) and subjected to immunofluorescent analysis with anti-PrimPol (green), anti-RPA2 (red), and counterstaining with DAPI (blue). Images taken with a 60x objective and deconvolved. (**B**) Cells were separated into Triton X-100 soluble and insoluble fractions and analysed by Western blot with the antibodies indicated. PrimPol^{HA} was detected with an anti-PrimPol antibody.

that HU treatment did not cause up-regulation of PrimPol protein, but rather PrimPol re-localised from the soluble to the insoluble fraction (Figure 6.13b), presumably interacting with DNA. These data suggest PrimPol may not only play a role in the tolerance of UV-C induced DNA damage, but could also function during nucleotide deprivation, this is discussed further below.

6.7. Discussion

Exposure of cells to UV-C light results in the formation of DNA damage. Covalent linkages form between adjacent pyrimidines that distort the DNA double helix, and if not repaired, pose a severe block to the DNA replication and transcription machineries. PrimPol, a novel primase-polymerase, re-localises within human cells following UV-C irradiation, tightly associating with chromatin at a number of discrete foci. These foci occur in cells undergoing DNA replication and may represent replication forks stalled at UV photoproducts. In line with this, human cells depleted of PrimPol contain more chromatin-bound RPA and an over-activated intra-S checkpoint following UV-C irradiation, indicative of increased replication fork stalling. These data are consistent with PrimPol functioning in the tolerance of UV-C induced DNA damage. An initial hypothesis was that PrimPol was required for re-initiating DNA synthesis on the leading strand downstream of the DNA lesion using its primase activity, a phenomenon demonstrated in *E. coli* (Heller and Marians, 2006; Yeeles and Marians, 2012) and suggested to occur in eukaryotes (Lehmann, 1972; Lopes *et al*, 2006, Elvers *et al*, 2011). However, work in the Doherty lab has demonstrated that human PrimPol, like the two diverse PrimPol homologues in Chapter 4, could catalyse TLS of UV photoproducts *in vitro*. PrimPol could extend from mismatched termini opposite a CPD and more strikingly, bypass a 6-4 photoproduct, which it did in an error-prone manner (Stanislaw Jozwiakowski and Aidan Doherty, unpublished; Chapter 4). Thus two mechanisms of PrimPol-dependant bypass of UV photoproducts can be envisaged, one in which PrimPol re-primes DNA synthesis downstream of the DNA lesion and a second in which PrimPol performs TLS of the lesion (Figure 6.14).

6.9.1. UV-induced foci formation and chromatin association of PrimPol

The assembly of proteins into discrete foci on chromatin following exposure to UV-C radiation is a common feature among numerous DNA damage signalling, repair, and tolerance enzymes, including TLS polymerases (Kannouche *et al*., 2001, 2003; Tissier *et al*., 2004; Ogi *et al.*, 2005; Andersen *et al*., 2011). Foci often represent locations of UV-C induced DNA damage, and in the case of Y-family polymerases, represent replication factories where DNA replication has stalled at a UV photoproduct (Kannouche *et al*., 2001, 2003; Tissier *et al*., 2004; Ogi *et al*., 2005). This remains to be determined for PrimPol; however, focal accumulations of PrimPol do represent a substantial number of PrimPol molecules concentrated into a number of discrete locations, which most likely correspond to ‘active’ PrimPol. This is supported by the correlation of DNA damage induced PrimPol foci and the sensitivities observed for PrimPol$^{-/-}$ chicken cells to particular DNA damaging agents (Julie Bianchi and Aidan Doherty, unpublished). PrimPol

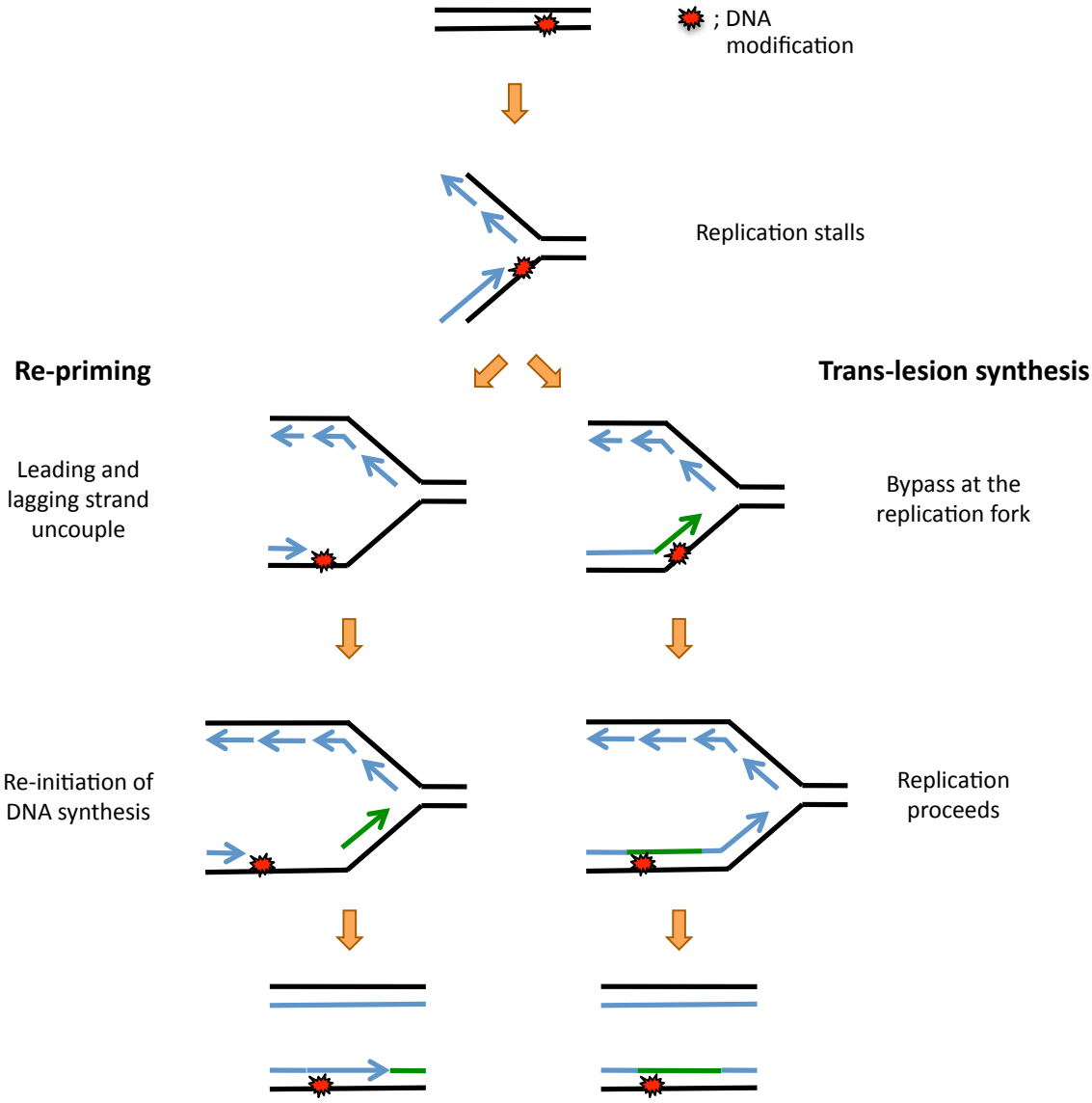


Figure 6.14. **Possible models of PrimPol-dependant DNA damage tolerance.**
DNA can become distorted and its bases modified by various environmental insults or endogenous processes, and if not corrected prior to DNA replication, can result in the physical blockage of replicative DNA polymerases. A DNA lesion on the leading strand is shown that causes functional uncoupling of leading and lagging strand DNA synthesis. This will generate a long stretch of unwound single-stranded DNA and PrimPol may re-prime DNA synthesis downstream of the lesion leaving a post-replication gap than can be subsequently filled by a trans-lesion synthesis polymerase. Or, PrimPol may perform trans-lesion DNA synthesis at the replication fork when the fork stalls, allowing continuation of DNA replication.

foci were observed following treatment of cells with agents known to cause replication fork stalling, such as UV-C radiation and HU, and further work in the Doherty lab showed both the “UV mimetic” 4-nitroquinoline 1-oxide (4NQO) and the alkylating agent MMS also triggered focal accumulation of PrimPol (Julie Bianchi and Aidan Doherty, unpublished). Furthermore, PrimPol foci were not observed following ionising radiation, which does not induce replication stalling and actually prevents DNA synthesis by inhibiting DNA initiation (Lamb *et al*., 1989). UV-dependant PrimPol foci occurred in ~35 % of cells at higher fluencies and these likely corresponded to S-phase cells, additionally PrimPol foci partially co-localised with detergent-resistant focal RPA. Although RPA is an essential component of NER (Araujo *et al*., 2000), detergent-resistant RPA foci largely represent RPA-coated single-stranded DNA present at stalled replication forks and not the much smaller RPA complexes formed at NER intermediates (Despras *et al*., 2010; Pathania *et al*., 2011; Diamant *et al*., 2012). Together, these data would be consistent with PrimPol associating with chromatin at sites of replication fork stalling following exposure of cells to UV-C irradiation. Future studies could determine whether PrimPol foci represent *bona fide* replication factories, as with Y-family polymerases, by analysing co-localisation with PCNA or DNA replication foci. In addition, mapping of the PrimPol domain(s) required for focus formation would prove insightful, and testing whether focus formation is dependant upon other proteins, such as Rad6/Rad18 dependant mono-ubiquitination of PCNA or interaction with Y-family polymerases such as REV1.

6.9.2. PrimPol mediated DNA damage tolerance in normal and XP-V cells

RNAi-mediated depletion of PrimPol in normal human fibroblasts led to increased chromatin-bound RPA following UV-C irradiation, consistent with more replication fork stalling. This is in line with analysis of spread DNA fibres from UV-C irradiated PrimPol$^{-/-}$ chicken cells, which demonstrated that PrimPol was required for replication of UV damaged DNA *in vivo* (Julie Bianchi and Aidan Doherty, unpublished). UV-C irradiation of PrimPol depleted fibroblasts also led to over-activation of the intra-S checkpoint kinase Chk1, which is consistent with PrimPol functioning in the tolerance of UV-C induced DNA damage. XP-V cells lacking functional Pol η, required to accurately bypass UV photoproducts, show enhanced activation of Chk1 following UV-C irradiation (Bomgarden *et al*., 2006 Despras *et al*., 2010). Similarly, cells depleted of Pol κ, which is known to bypass non-UV bulky adducts, were reported to have over-activation of Chk1 following treatment with benzo(α)pyrene-dihydrodiol epoxide (Bi *et al*., 2005). A defect in NER, however, which is required to remove UV photoproducts from DNA, results in no further activation of the intra-S checkpoint following UV-C irradiation (Bomgarden *et al*., 2006). Although the over-activation of the intra-S checkpoint observed in PrimPol

depleted cells could be accounted for by PrimPol RNAi increasing the proportion of S-phase cells, various human tumour cell lines treated with PrimPol RNAi show no significant pertubation of the cell cycle (Julie Bianchi, Laura Bailey, Aidan Doherty, unpublished). Further, cultures of PrimPol$^{-/-}$ chicken cells actually show fewer S-phase cells than wild-type cultures (Julie Bianchi and Aidan Doherty, unpublished).

Despite PrimPol being required for the tolerance of UV-C induced DNA damage, depletion of PrimPol did not sensitise normal human fibroblasts to UV-C induced killing. This may not be surprising; when considering PrimPol as a TLS polymerase, CPDs are efficiently bypassed by Pol η (McCulloch *et al*., 2004) whilst 6-4 photoproducts, accounting for ~30 % of UV-C induced lesions are rapidly repaired by NER (Mitchell *et al*., 1990). When considering PrimPol as a primase required for re-initiation of DNA synthesis downstream of a lesion, it would presumably only be required for leading strand, (and so ~50 % of) damage tolerance, as Okazaki fragment synthesis by Pol α-Prim would initiate downstream of lagging strand lesions, as observed *in vitro* (Svoboda and Vos, 1995). PrimPol depletion only rendered cells sensitive to UV-C induced killing in the absence of functional Pol η in XP-V cells. This demonstrates PrimPol operates in a DNA damage tolerance pathway that is non-epistatic with Pol η, and that loss of both these pathways lowers the threshold level of UV photoproducts that cells can tolerate. This may be due to the complementary TLS activities of these two polymerases: whilst Pol η can bypass CPDs and insert opposite 6-4 photoproducts, PrimPol can bypass 6-4 photoproducts and extend from CPDs (Johnson *et al*., 1999 and 2001). Future studies could determine PrimPol's contribution to TLS *in vivo*, using plasmids containing UV photoproducts (Ziv *et al*., 2012) or antibodies specific to the lesions (Temviriyanukul *et al*., 2012). Determining whether PrimPol contributes to UV-C induced mutagenesis will also prove insightful, especially for XP-V patients, in which mutagenesis leads to carcinogenesis.

6.9.3. Non UV-induced DNA damage tolerance roles of PrimPol

In addition to PrimPol's damage tolerance role following UV-C irradiation, nucleotide deprivation by HU treatment, which causes replication fork stalling but in the absence of DNA lesions, resulted in PrimPol's association with chromatin at discrete foci. This is also a common feature of Y-family polymerases (Ogi *et al*., 2005; de Feraudy *et al*., 2007), but a role for these enzymes during nucleotide deprivation is not yet fully understood. It is conceivable that foci formation following HU treatment could represent the polymerases presence at replication factories that are now stalled. Although it has been demonstrated that Pol η and the *E. coli* Y-family polymerases Pol IV and Pol V are

required for DNA replication during HU treatments (de Feraudy *et al*., 2007; Godoy *et al*., 2006). In the case of the latter, it was suggested that because these Y-family polymerases have much lower K_ms for dNTPs than replicative polymerases, they could take over DNA replication at times of nucleotide deprivation to prevent replication fork collapse (Godoy *et al*., 2006). In addition to the roles proposed for Y-family polymerases, it is also conceivable that PrimPol may be attempting to re-initiate DNA synthesis downstream of the stalled fork through its primase activity. Further studies could elucidate whether PrimPol is required for DNA replication or restart following HU treatment.

PrimPol depletion in normal human fibroblasts led to a significant growth detect, indicating a role for PrimPol in normal cell proliferation. Further work in Aidan Doherty's laboratory demonstrated that this proliferative defect was independent of PrimPol's mitochondrial function using cells depleted of their mitochondrial DNA (Laura Bailey and Aidan Doherty, unpublished), suggesting that it is PrimPol's nuclear role that is responsible. It is also intriguing that PrimPol depletion in XP-V cells resulted in a greater proliferative defect. Consistent with a role for PrimPol in an unperturbed S-phase, analysis of spread DNA fibres from PrimPol$^{-/-}$ chicken cells showed that overall replication fork speeds were reduced in these cells (Julie Bianchi and Aidan Doherty, unpublished). A role for Pol η during unperturbed DNA replication has been described, showing that it is required for replication of chromosomal fragile sites (Rey *et al*., 2009). Many endogenous processes can result in DNA modifications, whether DNA lesions or structures, which impede replication, and it is possible that PrimPol is involved in the tolerance of these modifications.

6.9.4. PrimPol – a TLS polymerase or primase, or perhaps both?

One important question underlying PrimPol's role in damage tolerance is whether PrimPol functions as a TLS polymerase, or a primase, or perhaps both (Figure 6.14). Re-initiation of DNA synthesis downstream of a DNA lesion was demonstrated in *E. coli* to be an inherent property of the replisome, being dependant on the replicative DnaG primase (Yeeles and Marians, 2012). Similarly, electron micrographs of DNA replicated in UV irradiated *S. cerevisiae* show single-stranded gaps present behind replication forks, consistent with re-priming events (Lopes *et al*., 2006). Yeast, however, lack an identifiable PrimPol homologue, suggesting Pol α-Prim is responsible. Re-priming occurs on the lagging strand of DNA replication by Pol α-Prim, and is inherently DNA damage tolerant due to the discontinuous nature of Okazaki fragment synthesis (Svoboda and Vos, 1995). Pol α-Prim is therefore already present at the replication fork, and upon the

replisome encountering a DNA lesion on the leading strand and the uncoupling of leading and lagging strand synthesis occurs, it would seem likely that Pol α-Prim would be responsible for this leading strand re-initiation on the unwound single-stranded DNA. It could be speculated that PrimPol may provide a more flexible option perhaps, being capable of a number of nucleotidyl transferase activities including TLS, and so may be required for re-initiation on particularly challenging templates containing clustered DNA damage. But yet again, presumably DNA synthesis would re-initiate downstream of these difficult templates.

A further point is that primed single-stranded DNA is required for activation of the intra-S checkpoint (MacDougall *et al*., 2007). The accumulation of newly synthesised primers at stalled replication forks, shown to require Pol α-Prim, and their elongation by the leading and lagging strand polymerases contributes to the activation of Chk1 in *Xenopus* egg extracts (Van *et al*., 2010). Therefore, if PrimPol was required to synthesise these primers at stalled replication forks, in an attempt to re-initiate DNA synthesis, this would be expected to result in reduced Chk1 activation in UV irradiated PrimPol depleted cells. However, this was not the case, as an increase in Chk1 activation was observed. Additionally, PrimPol$^{-/-}$ chicken cells were just as capable as wild-type cells in converting newly synthesised small molecular weight DNA into large molecular weight DNA following UV-C irradiation (Julie Bianchi and Aidan Doherty, unpublished). This is consistent with these cells having no defect in filling single-stranded DNA gaps that encompass UV photoproducts, which are created by re-priming events occurring downstream of the DNA lesion. This again, suggests PrimPol is not required for re-priming downstream of a replication blocking lesion. On the other hand, evidence is yet to be obtained of PrimPol-dependant TLS of UV photoproducts *in vivo*. PrimPol preferentially incorporates T opposite the 3' T of a T-T 6-4 photoproduct, and C or G opposite the 5' T (Chapter 4; Stanislaw Jozwiakowski and Aidan Doherty, unpublished). Bypass of 6-4 photoproducts *in vivo* has been reported to be largely error-free, although the most frequent mis-insertion was reported to be T opposite the 3' T (Yoon *et al*., 2010a; Szuts *et al*., 2008; Hendel *et al*., 2008, Hirota *et al*., 2010), which could be PrimPol-dependant. Further experiments are required to clarify the *in vivo* role of PrimPol in the tolerance of UV-C induced DNA damage.

Chapter 7

Towards the identification of human PrimPol protein partners

7.1. Introduction

Protein-protein interactions play a fundamental role in biological processes. Among the plethora of examples are the network of proteins required to propagate the eukaryotic replication fork (Johnson and O'Donnell, 2005) and the protein complexes responsible for repairing DNA damage (Hoeijmakers, 2001). Protein-protein interactions often play important roles in regulating the activity of enzymes, and this is true of both DNA polymerases and primases (Kelman *et al*., 1998; Frick and Richardson, 2001). PrimPol is a novel primase-polymerase present in the nucleus and mitochondrion of human cells (Chapter 3). Before the discovery of PrimPol's role in DNA damage tolerance (Chapters 4 and 6) initial attempts to determine the function of this novel enzyme focused on identifying its protein partners, which was the aim of this current chapter. The rationale was that identifying proteins that interact with PrimPol would highlight the cellular pathways in which PrimPol operates, and possibly provide a mechanistic insight into the role of this novel enzyme. A number of approaches were adopted to identify PrimPol protein partners, with the central method being the affinity purification of PrimPol from cultured human cells and analysis of the purifications by mass-spectrometry.

7.2. Identification of *in vivo* protein partners using the *Strep*-tag system

The primary strategy to identify proteins that interact with PrimPol was to purify PrimPol from cultured human cells and identify any co-purifying proteins using mass-spectrometry. To facilitate affinity purification of PrimPol it was decided to express PrimPol fused to a *Strep*-tag, which exploits the high affinity and specific binding between streptavidin and its natural ligand biotin (Schmidt *et al*., 1996). Specifically, the eight amino acid long *Strep*-Tag II (WSHPQFEK) was used, which allows affinity purification with the streptavidin derivative *Strep*-Tactin and specific elution with desthiobiotin (Schmidt *et al*., 1996; Voss and Skerra, 1997; Korndörfer and Skerra, 2002). Affinity purified *Strep*-tagged PrimPol and co-purifying proteins could then be resolved by SDS-PAGE and analysed by mass-spectrometry (Figure 7.1).

7.2.1. Over-expressed PrimPol was initially degraded

The Flp-In T-REx system was used to generate a stable cell line with inducible expression of *Strep*-tagged PrimPol. The PrimPol cDNA was cloned without a stop codon in frame with tandem Flag and *Strep*-Tag II epitopes into an inducible expression vector, generating the construct pcDNA5/FRT/TO:PrimPol-Flag-*Strep*-Tag II. This was then used to make a stable inducible cell line in HEK-293 cell derivates (Flp-In T-REx-293

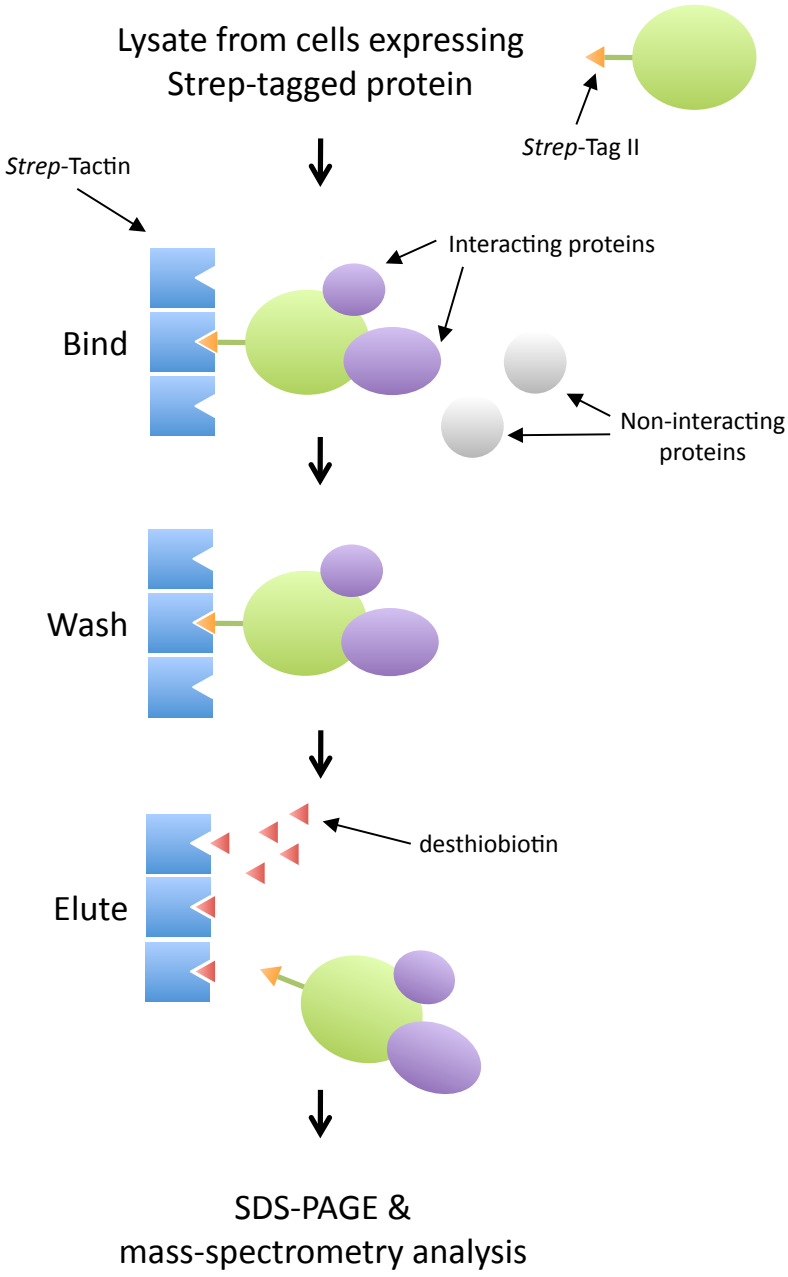


Figure 7.1. Schematic overview of *Strep*-tag affinity purification method.
The protein of interest is expressed in human cells fused to a *Strep*-Tag II. Lysate is prepared from these cells and applied to *Strep*-TACTIN resin, which selectively binds the *Strep*-Tag II with high affinity. This immobilises the protein fused to the *Strep*-Tag II and associated protein complexes, whilst non-interacting proteins will be washed away. Specifically bound proteins can then be eluted using desthiobiotin and released protein complexes resolved by SDS-PAGE and analysed by mass-spectrometry.

cells) (described in section 5.2.1), generating a cell line with doxycycline-inducible expression of PrimPol fused to a carboxyl-terminal Flag and *Strep*-Tag II (PrimPol^{FlagStrep}).

Although these stable, inducible cell lines eventually worked as desired (Chapters 5 and 6), they were initially problematic, and this affected the affinity purification process. Western blot analysis of lysates prepared from cells grown in the presence of doxycycline revealed that over-expressed PrimPol was largely degraded, with very little full-length PrimPol being detected (Figure 7.2a and b). Further, no endogenous PrimPol was detected (Figure 7.2a and b). This is possibly due to leaky expression of recombinant PrimPol due to trace amount of tetracycline in the cell media, which is having a dominant negative affect resulting in the degradation of endogenous PrimPol. Cell lines were remade selecting both clonal and pooled populations, and additional cell lines expressing HA-tagged PrimPol were generated, but the same result was observed (not shown). Expression conditions were optimised with this caveat, and cells grown in the presence of 10 ng/ml doxycycline for 24 hours gave the most promising expression (Figure 7.2a and b). Despite the high level of degradation observed in whole cell lysates, Western blot analysis of mitochondrial lysates revealed a significant amount of full-length over-expressed PrimPol, with no detectable degradation products (Figure 7.2c), and so it was decided to proceed with the affinity purification.

7.2.2 Affinity purification of *Strep*-tagged PrimPol from whole cell lysate

As PrimPol is localised to both nuclei and mitochondria, the goal was to perform affinity purification from these two compartments, but first, a preliminary affinity purification from whole cell extract was attempted. Flp-In T-REx-293 cells engineered for inducible expression of *Strep*-tagged PrimPol were grown in the presence or absence of 10 ng/ml doxycycline for ~24 hours and then harvested, with a ~1 g cell pellet obtained for both the induced and uninduced cultures. The soluble whole cell lysate was prepared and *Strep*-Tactin resin added. The mixture was then applied to a gravity flow column and washed extensively before successive elutions with desthiobiotin. Western blot analysis shows that full-length over-expressed PrimPol and various truncations were enriched in the elutions from the induced sample, both in comparison to the input, but also to the non-induced sample, indicating affinity purification of PrimPol was successful (Figure 7.3a). Elutions 2 and 3 were selected from the induced and non-induced and resolved on a gradient SDS-polyacrylamide gel and stained with colloidal Coomassie (Figure 7.3b), and although no obvious difference was visible, it was decided to analyse these purification by mass-spectrometry. Whole lane gel extraction was performed and samples were analysed in collaboration with Mark Skehel (CRUK, Clare Hall).

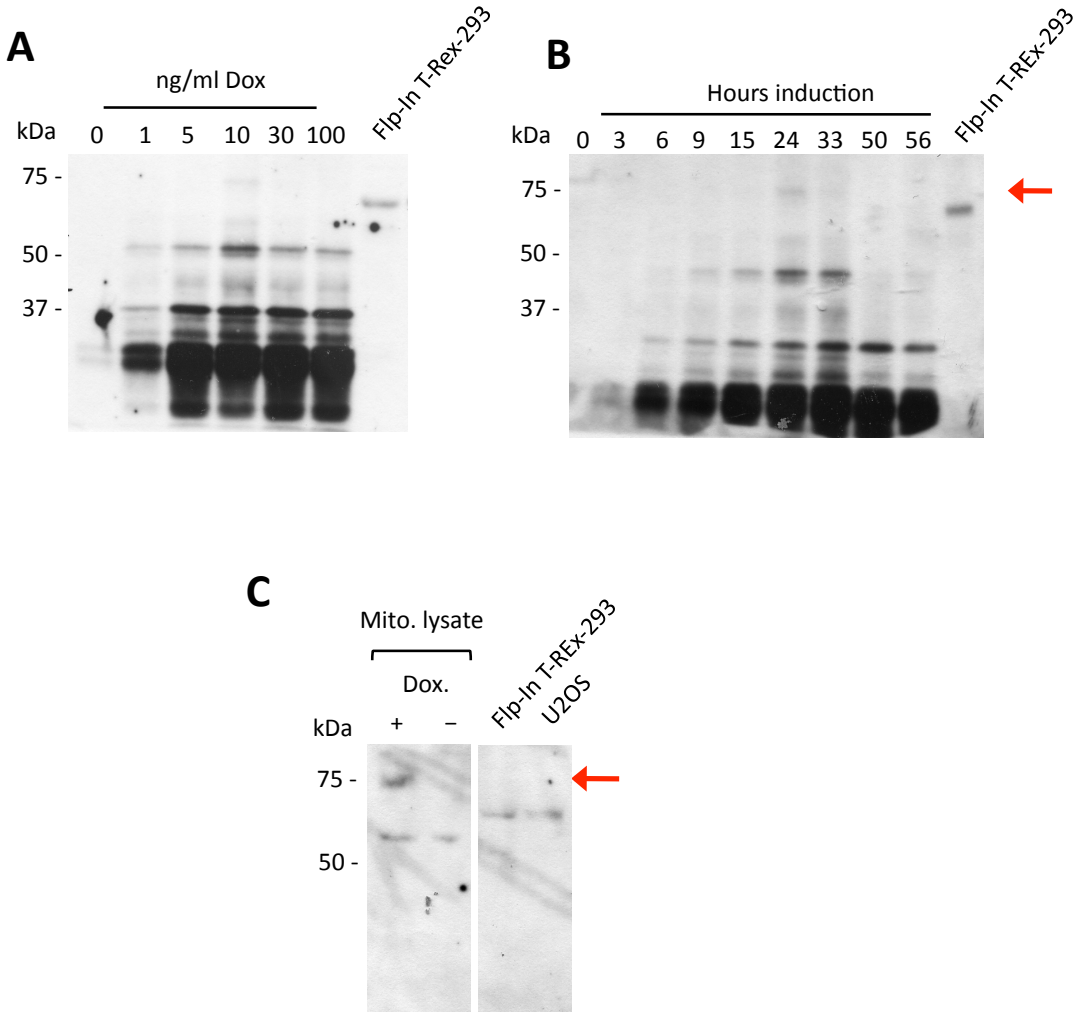


Figure 7.2. Stably over-expressed PrimPol was initially degraded in human cells, however only full-length was imported into mitochondria. Flp-In T-REx-293 cells engineered for inducible expression of PrimPol^{FlagStrep} were induced by addition of indicated concentrations of doxycycline for ~24 hours (**A**) or 10 ng/ml doxycycline for the times indicated (**B**), and whole cell extracts analysed by Western blot with an anti-PrimPol antibody. (**C**) Cells were grown in the presence or absence of 10 ng/ml doxycycline for ~24 hours and mitochondria isolated and analysed by Western blot with an anti-PrimPol antibody. Flp-In T-REx-293 refers to the parental cell line. Red arrow indicates full-length PrimPol^{FlagStrep}.

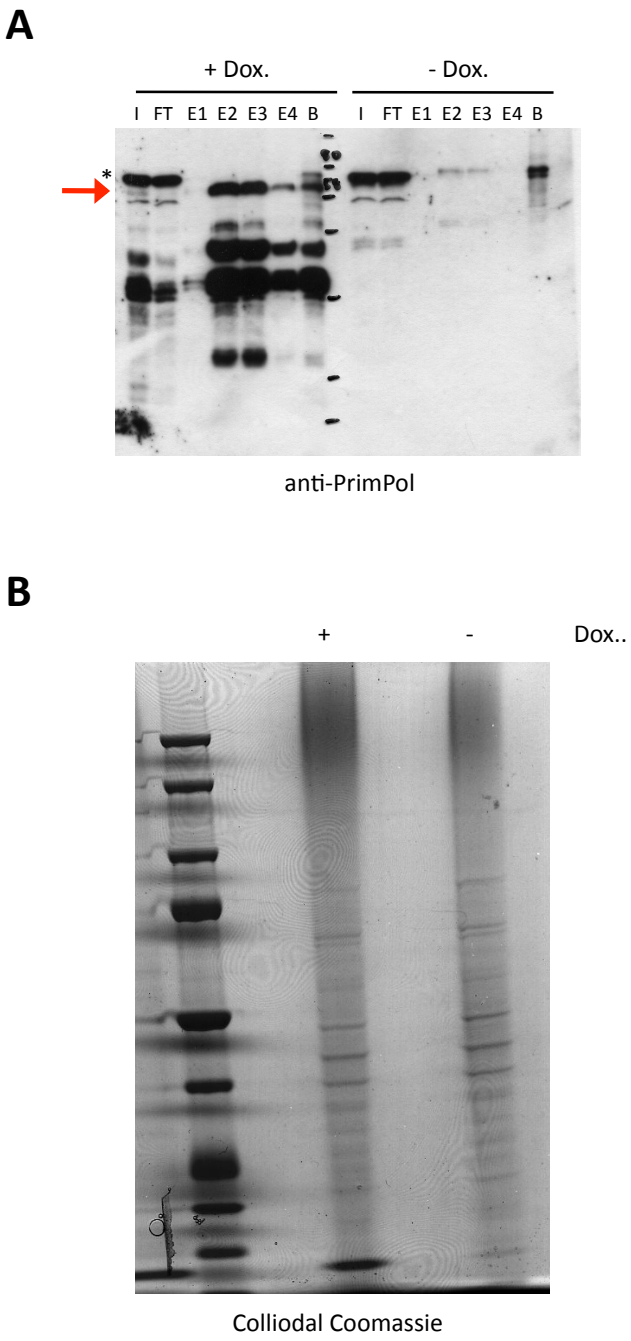


Figure 7.3. Affinity purification of *Strep*-tagged PrimPol from whole cell lysate.
(**A**) Large cultures of Flp-In T-REx-293 cells engineered for inducible expression of PrimPol^{FlagStrep} were grown in the presence or absence of doxycycline (10 ng/ml, 16-24 hours) and then PrimPol^{FlagStrep} was affinity purified using *Strep*-Tactin resin from the soluble cell lysate. Fractions from the affinity purification were analysed by Western blot with an anti-PrimPol antibody: input (I), flow-through (FT), elutions (E), *Strep*-Tactin beads (B). Red arrow indicates full-length PrimPol^{FlagStrep}, asterisk indicates non-specific band. (**B**) Elutions 2 and 3 from the induced and non-induced cultures were pooled and resolved by SDS-PAGE on a 4-20 % gel and stained with colloidal Coomassie. Each lane was excised in ~2 mm gel slices and prepared for mass-spectrometry analysis.

7.2.3. Mass-spectrometry analysis of *Strep*-tagged PrimPol purifications

The proteins identified in the mass-spectrometry analysis were ranked according to percentage of total spectra in the induced sample, and the fold enrichment calculated for each. A large set of proteins (1249) was identified, of these ~550 were present only in the induced sample and a further 65 showed a 3 or more-fold enrichment, with PrimPol having a 20-fold enrichment. Input of these proteins into the Database of Annotation, Visualisation, and Integrated Discovery (DAVID) (Dennis *et al*., 2003; Huang *et al*., 2008) clustered these proteins into a number of functionally related groups (Figure 7.4). Consistent with the dual localisation of PrimPol, two of the predominant groups were nuclear and mitochondrial proteins. A large proportion of DNA and nucleotide binding proteins were also present, and more specifically some proteins involved in DNA replication and repair, such as some RPA and MCM subunits. In contrast, no mitochondrial replication enzymes were present, although mtSSB was present below the 3-fold enrichment cut-off. Core mitochondrial nucleoid components were present, such as ATAD3A and PDIP2. One ominous hit was the 22-fold enrichment of streptavidin in the induced sample, suggesting contamination of this sample with *Strep*-Tactin resin. The mass-spectrometry data was preliminary and so in-depth analysis would be superfluous, further experiments need to be performed with more stringency to increase the likelihood of identifying a PrimPol protein partner.

7.2.4 Validating potential protein partners - RPA and mtSSB co-purify with PrimPol from human cells

To validate the potential PrimPol interacting proteins from the preliminary mass-spectrometry analysis, small-scale affinity purifications of *Strep*-tagged PrimPol from whole cell lysate were performed and analysed by Western blot. It was at this time that it became apparent that over-expressed PrimPol was no longer being degraded. Following addition of doxycycline a predominant species of ~69 kDa was detected by Western blot analysis with an anti-PrimPol antibody (Figure 7.5), and furthermore, endogenous PrimPol was also detected (Figure 7.5). These cells were the same as previously used when PrimPol degradation was visible, except these cells were a higher passage. *Strep*-tagged PrimPol was largely depleted from the soluble cell lysate following addition of the *Strep*-Tactin resin, and was successfully eluted using desthiobiotin (Figure 7.5). Very little *Strep*-tagged PrimPol was visible in the non-induced lysate, although leaky expression was observed (Figure 7.5). Analysis of the affinity purification with antibodies for RPA subunits 1 and 2, and the mitochondrial equivalent mtSSB, all gave specific bands in the elutions (Figure 7.5), suggesting PrimPol may associate with these proteins. ATAD3, a

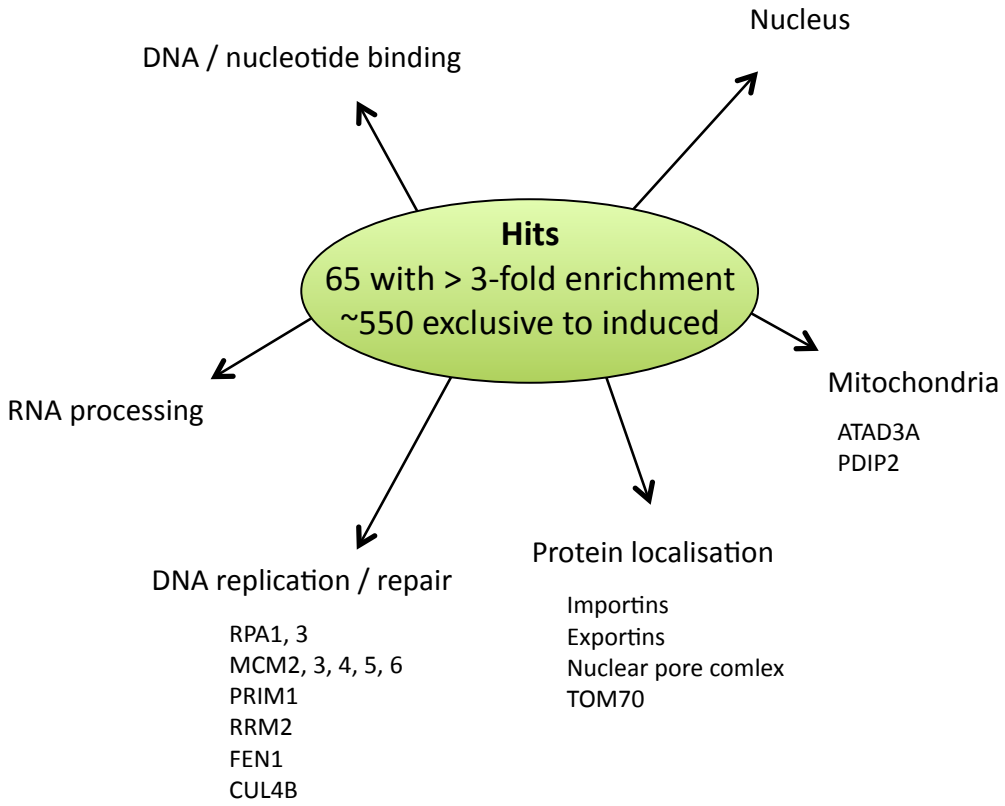


Figure 7.4. Summary of proteins identified in mass-spectrometry analysis.
Proteins identified by mass-spectrometry to be co-purifying with *Strep*-tagged PrimPol were clustered into functionally related groups using DAVID (Dennis *et al*, 2003; Huang *et al*, 2008). Groups identified are displayed along with some examples of proteins within that group.

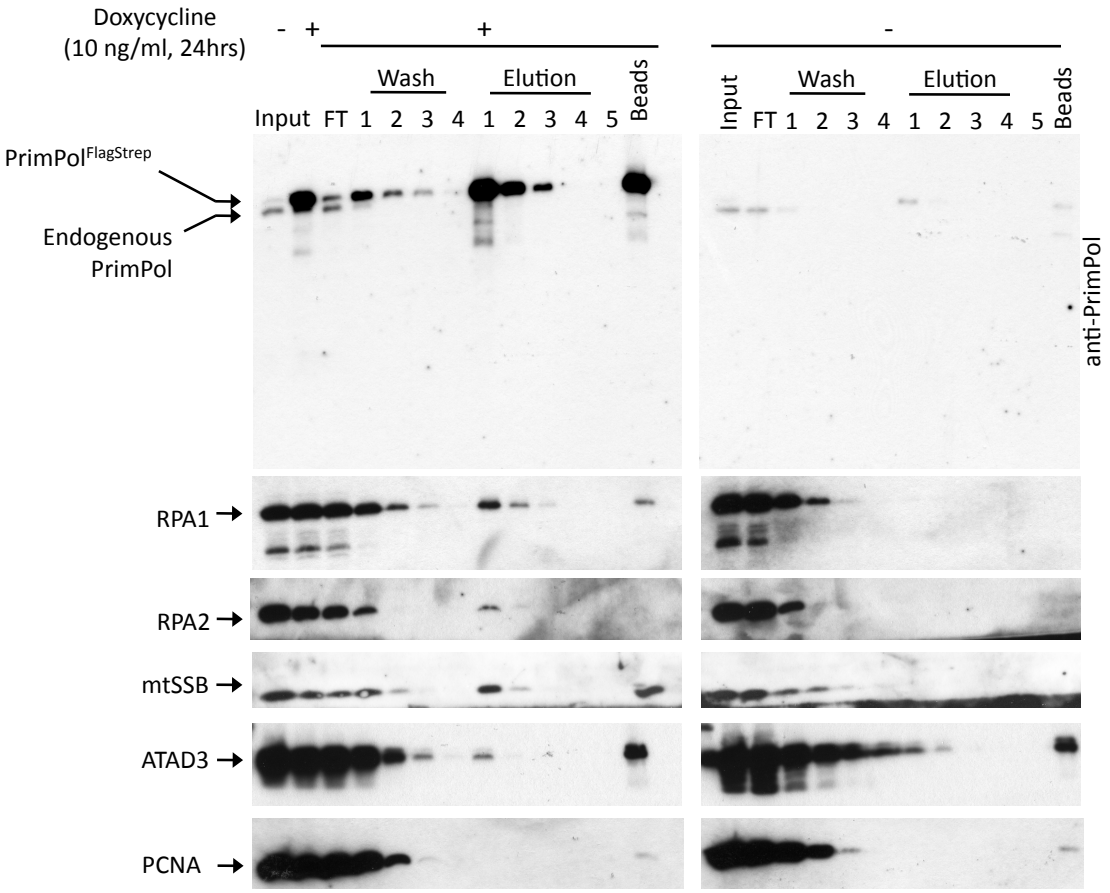


Figure 7.5. Validating mass-spectrometry hits - PrimPol co-purifies with mtSSB and RPA.
Flp-In T-REx-293 cells engineered for inducible expression of PrimPol^{FlagStrep} were grown in the presence or absence of doxycycline (10 ng/ml, 24 hours) and PrimPol^{FlagStrep} was affinty purified from the soluble cell lysate using *Strep*-Tactin resin. Fractions from the affinity purification were analysed by Western blot with the antibodies indicated.

mitochondrial membrane associated ATPase and core nucleoid component, was also detected in the elutions, however ATAD3 did appear to bind to the *Strep*-tactin resin in the non-induced sample (Figure 7.5). Other proteins were probed for, such as MCM2, PCNA, and PARP, but these were not detected in the elutions. In addition, proteins that were not identified in the mass-spectrometry analysis but were likely candidates for interaction, such as the mitochondrial replicase Pol γ and helicase Twinkle, were probed for but again no co-elution was observed (Figure 7.5 and data not shown). Although the preliminary hits obtained from the mass-spectrometry were many, the small-scale purifications have validated some of them as proteins that co-purify with PrimPol, such as the single-stranded DNA binding proteins RPA and mtSSB. It remains to be determined whether these interactions are protein-protein or DNA-mediated.

7.2.5. PrimPol is phosphorylated *in vivo* on serine 499 and 501

In addition to identifying potential protein partners, mass-spectrometry can also be used to identify post-translational modifications. In the early stages of PrimPol characterisation it was noted that PrimPol resolved as a doublet by SDS-PAGE. Therefore, to test whether this doublet was due to the common post-translational modification of phosphorlyation, *Strep*-tagged PrimPol was affinity purified and treated with lambda phosphatase in the presence or absence of a phosphatase inhibitor cocktail. Affinity purified PrimPol protein resolved as a doublet and addition of lamda phosphatase led to a significant reduction of the upper band, which was specific on phosphatase activity, indicating that PrimPol was phosphorylated *in vivo* (Figure 7.6a). Thus, during mass-spectrometry analysis of affinity purified *Strep*-tagged PrimPol (section 7.2.2), the phosphorylation site(s) was investigated. PrimPol was phosphorylated on serine residues 499 and 501 located at the carboxyl-terminus of the protein (Figure 7.6b). Further investigation into the significance of these phosphorylations will no doubt impact upon our understanding of this novel enzyme.

7.3. Identification of potential protein partners *in vitro* using recombinant candidate proteins

In addition to the affinity purification and mass-spectrometry strategy, various other approaches were tried to identify PrimPol protein partners. One of which was a candidate approach in which proteins were selected that are likely to interact with PrimPol and the interaction tested *in vitro* using far-Western slot blot analysis with recombinant proteins. Components of the mitochondrial replisome were likely candidates and the mtDNA replicase Pol γ was first tested.

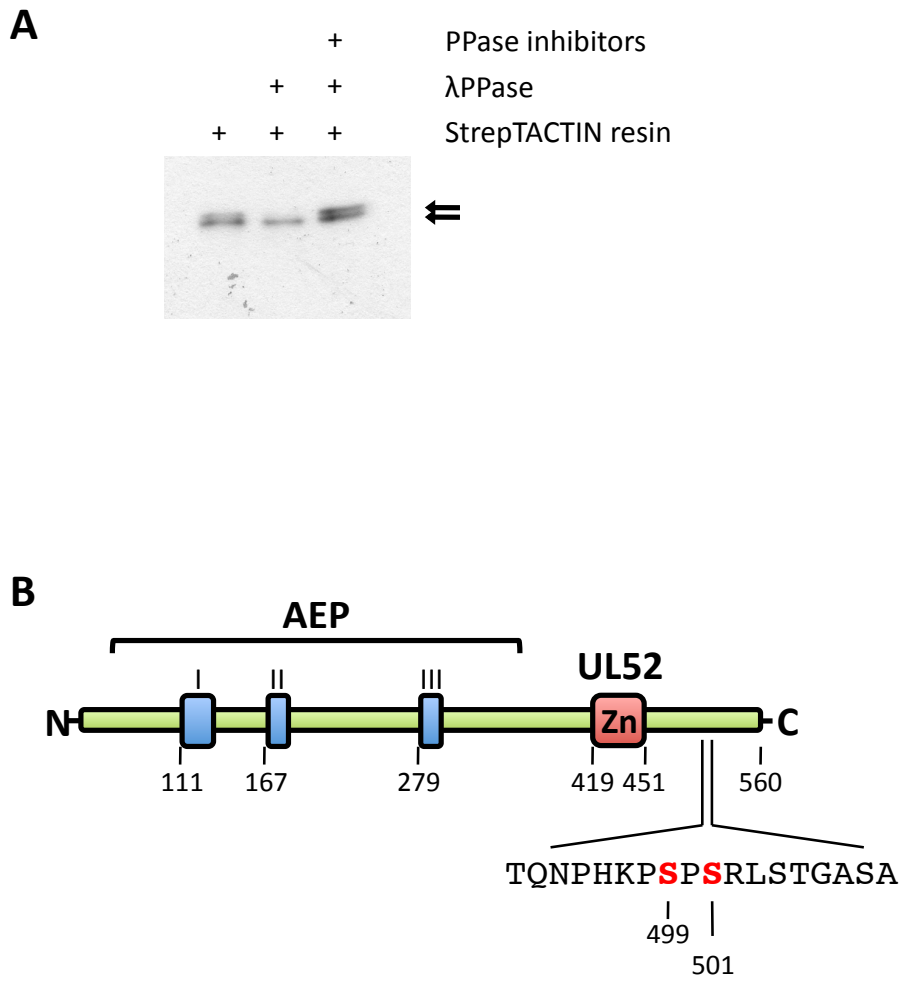


Figure 7.6. PrimPol is phosphorylated *in vivo*.
(**A**) Flp-In T-REx-293 cells engineered for inducible expression of PrimPol^{FlagStrep} were induced for expression (10 ng/ml doxycycline, 16-24 hours) and PrimPol^{FlagStrep} was affinity purified from the cell lysate using *Strep*-Tactin resin. Purified PrimPol was then treated with Lambda Phosphatase in the presence or absence of a phosphatase inhibitor cocktail. Samples were then analysed by Western blot with an anti-PrimPol antibody. Arrows indicate PrimPol doublet. (**B**) Affinty purified PrimPol^{FlagStrep} was analysed by mass-spectrometry to detect phosphorylation. Phosphorylated serines are indicated in red.

7.3.1 Purification of recombinant PolG2 produced in *E. coli*

Pol γ exists as a heterotrimer of one catalytic subunit (PolG1) and two accessory subunits (PolG2) (Kaguni, 2004). Over-expression and purification of recombinant human PolG2 has been reported previously in *E. coli* and was shown to be a single-step procedure (Di Re *et al*., 2009), unlike the more complex preparation of the catalytic subunit (Lee *et al*., 2009). For this reason, initial efforts focused on the accessory PolG2 subunit. A culture of *E. coli* B834s transformed with an expression vector encoding human PolG2 with an amino-terminal 10-histidine tag (pRUN:POLG2, Table 2.5) (gift from Ian Holt, MRC, Cambridge) (Di Re *et al*., 2009) was grown until exponential phase and induced for expression by addition of 0.6 mM IPTG and grown overnight at 25 ºC. Soluble lysate was prepared from these cells and subjected to Ni^{2+}-NTA agarose affinity chromatography and the bound proteins eluted by increasing concentrations of imidazole. Recombinant human PolG2 was purified to homogeneity in this single-step, visible by SDS-PAGE analysis as a protein with a relative mass of ~50 kDa in the 300 mM imidazole elution (Figure 7.7).

7.3.2. PrimPol interacts with the mtDNA replicase Pol γ *in vitro*

Recombinant PolG2 was then used in Far-Western blot analysis. Increasing concentrations of recombinant human PolG2 and recombinant human PrimPol (gift from Julie Bianchi, Doherty Lab) (Figure 7.8a) were slot-blotted onto a membrane and probed with the other recombinant protein followed by the corresponding antibody. PolG2 could be detected using recombinant PrimPol and likewise, PrimPol could be detected using recombinant PolG2 (Figure 7.8b), confirming an *in vitro* interaction between these two proteins. This potential interaction between PrimPol and the mtDNA replicase was investigated further using recombinant human PolG1 (gift from Whitney Yin, University of Austin, Texas). Increasing concentrations of PolG1 and as a negative control the bacterial NHEJ polymerase PolDom (gift from Nigel Brissett, Doherty Lab), were blotted onto a membrane and probed with recombinant PrimPol. Whilst PrimPol did not detect recombinant PolDom, it did detect PolG1 (Figure 7.8c). Together these data suggest that PrimPol does interact with the mtDNA replicase, although further experiments need to be performed to determine whether this is the case *in vivo*.

7.4. Identification of protein partners using the Yeast-Two Hybrid assay

The yeast two-hybrid assay was also chosen as a method to identify PrimPol interacting proteins. This reporter-based assay exploits the modular structure of transcription

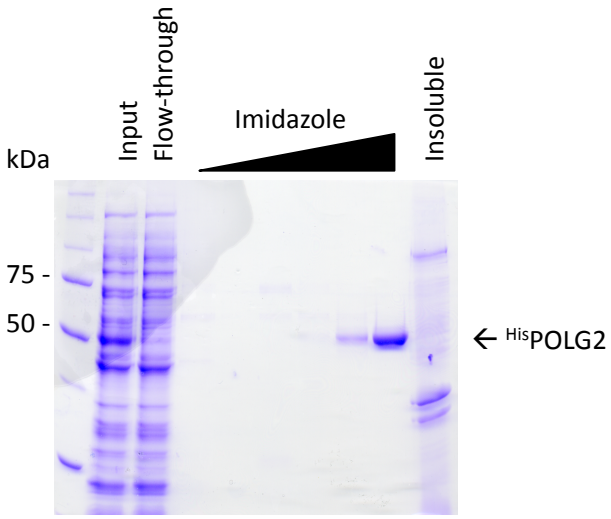


Figure 7.7. Purification of the human mtDNA replicase accessory subunit PolG2 from *E. coli*
A 2 L culture of B843s transformed with pRUN:PolG2 was induced for ^{His}PolG2 expression by addition of 0.6 mM IPTG and incubated at 25 ^{o}C overnight. The soluble cell lysate (prepared as detailed in Material and Methods except for addition of 0.2 % Triton X-100) was applied to a Ni^{2+}-NTA column and following washes, ^{His}PolG2 was eluted with 300 mM imidazole. Fractions from the purification were resolved by SDS-PAGE and Coomassie stained.

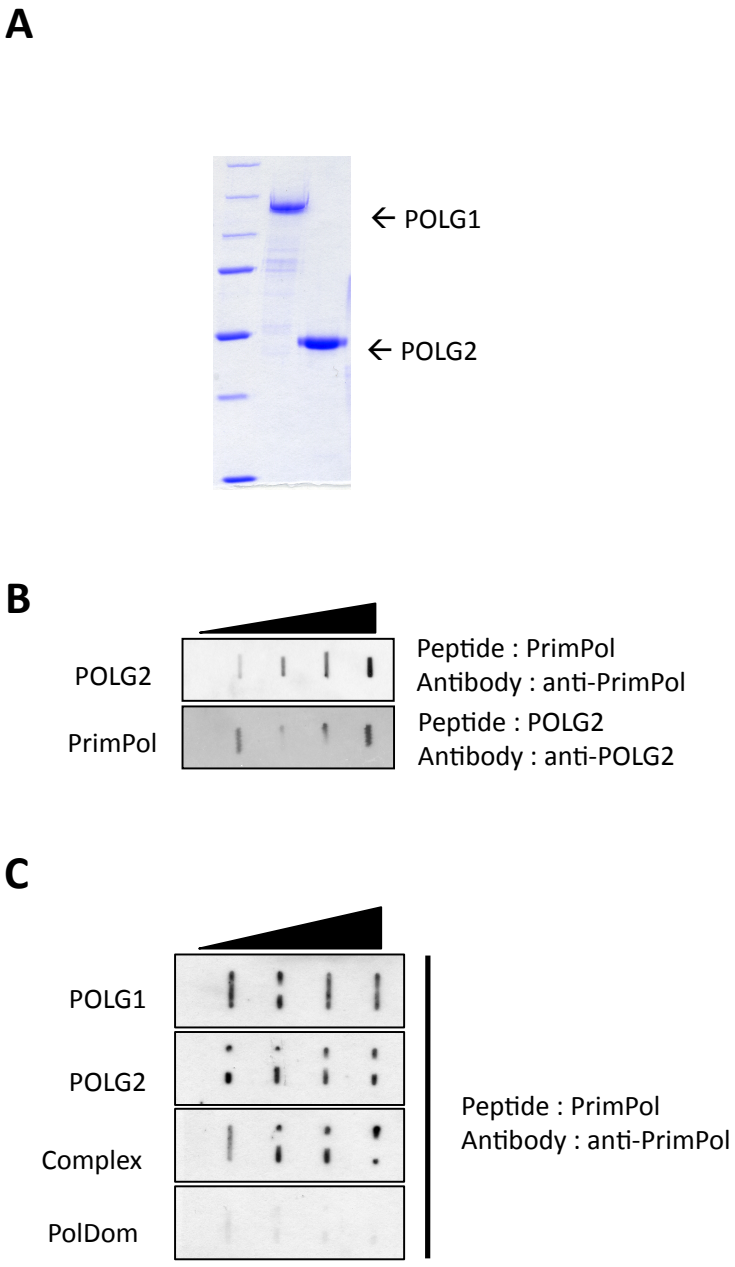


Figure 7.8. Evidence for PrimPol interaction with the mtDNA replicase Pol γ.
(**A**) Prepared PolG1 and 2 were resolved by SDS-PAGE and stained with Coomassie. (**B**) *In vitro* interaction of PrimPol and PolG2. Recombinant PrimPol and PolG2 were slot blotted onto a membrane (0, 0.25, 0.5, 1, and 2 µg) and probed with the protein (1 µg) and antibody indicated. (**C**) Slot blot experiment; recombinant PolG1, PolG2, and PolDom were blotted onto a membrane as described in (A) and were probed with recombinant PrimPol (1 µg). PrimPol was detected using an anti-PrimPol antibody.

factors, in particular the *S. cerevisiae* GAL4 transcriptional activator (Fields and Song, 1989). The DNA binding domain of the GAL4 protein is expressed as a fusion to protein X (bait), and the transcriptional activation domain is expressed as a fusion to protein Y (prey). Interaction of protein X and Y results in the recruitment of the basal transcriptional machinery to the promoters of a number of reporter genes, thus activating their transcription and providing evidence for protein-protein interaction (Fields and Song, 1989) (Figure 7.9). Initially the yeast-two hybrid assay was to be used in a candidate approach, using specific cDNAs in a small-scale study, but the eventual aim was to perform a genome-wide screen with a human cDNA library.

7.4.1. PrimPol carboxyl-terminal containing the UL52 zinc finger auto-activates

The PrimPol cDNA was cloned into the yeast-two hybrid vector pGBKT7, creating the construct pGBKT7:PrimPol (gift from Andrew Green, Doherty Lab) (Table 2.10). This construct encodes for a PrimPol fusion protein with the GAL4 DNA binding domain. It was first tested whether expression of PrimPol in the Y190 *S. cerevisiae* strain resulted in auto-activation of the *LacZ* and *HIS3* reporter genes; *LacZ* transcription results in expression of β-galactosidase and transcription of *HIS3* allows growth on media containing no histidine. Y190 were co-transformed with pGBKT7:PrimPol and empty pACT2, which encodes for the GAL4 transcriptional activation domain, and cells containing both plasmids tested for transcription of the reporter genes. Expression of the PrimPol-GAL4 binding domain fusion, in the absence of a prey protein, resulted in auto-activation of both reporter genes; transformants tested positive for β-galactosidase activity and were able to grow on media containing 3-AT, a competitive inhibitor of the yeast *HIS3* gene product (Durfee *et al*., 1993). Thus, expression of the PrimPol – DNA binding domain fusion results in auto-activation, and therefore this fusion protein cannot be used in a yeast two-hybrid assay.

To overcome this problem I truncated the PrimPol protein in order to remove the residues responsible for auto-activation. Two truncations were generated in total, both using PCR methods: one encompassing the AEP domain (1-341) and a second encompassing the UL52 domain (370-560) (Figure 7.10a). For expression of the AEP domain, complementary primers were designed to introduce a stop codon in the PrimPol cDNA truncating the protein at residue 341 (Table 2.11). These primers were used in a site-directed mutagenesis PCR reaction with the construct pGBKT7:PrimPol to generate the construct pGBKT7:PrimPol 1-341 (Table 2.10). In order to express the UL52 domain an inverse deletion PCR was performed. This method uses a primer pair in a “back-to-back” orientation (opposite to usual PCR) that contain a 5’ phosphate, which allows

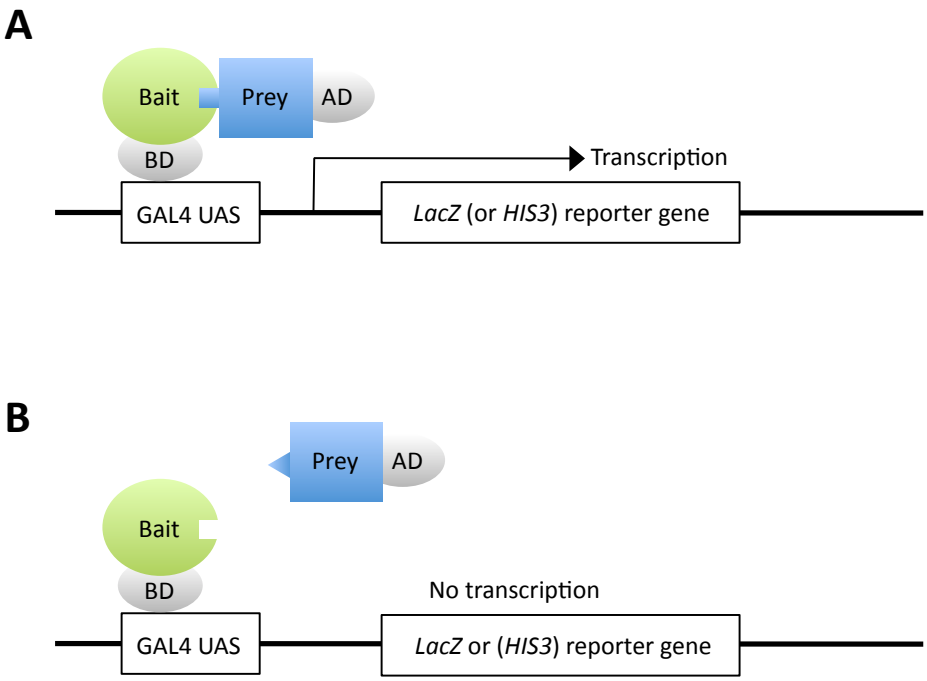


Figure 7.9. Schematic detailing the yeast two-hybrid assay.
(**A**) The fusion of the bait protein to the GAL4 DNA binding domain (BD) binds to the GAL1 upstream activating sequence (UAS). Interaction of the bait protein to the prey protein, which is fused to the GAL4 transcriptional activation domain (AD), reconstitutes the function of the GAL4 protein and allows expression of the reporter genes. (**B**) If no interaction of the bait and prey fusion occurs, the GAL4 transcription factor is not reconstituted and so the reporter genes are not expressed.

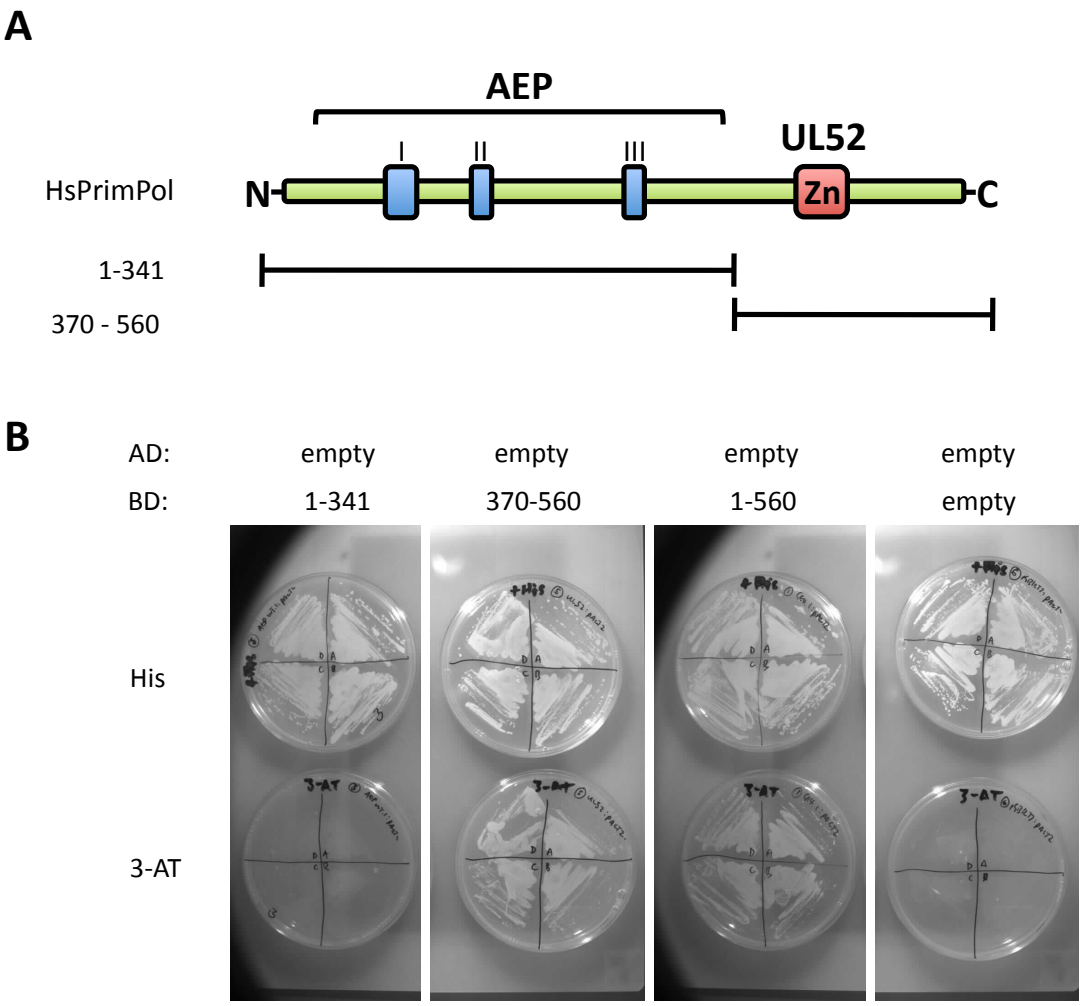


Figure 7.10. The PrimPol C-terminus (370-560), encompassing the UL52 zinc finger, is responsible for auto-activation in the yeast two-hybrid assay. (**A**) Schematic detailing PrimPol truncations used in yeast two-hybrid. (**B**) Full length PrimPol and truncations were expressed in yeast Y190 fused to the GAL4 DNA binding domain and auto-activation assessed by growth on plates containing 3-AT, a competitive inhibitor of the product of the *HIS3* gene.

ligation of the PCR product. A desired sequence can then be deleted by having the primers flank this region. Primers were designed (Table 2.11) and an inverse deletion PCR performed with pGBKT7-PrimPol in order to remove the sequence corresponding to residues 2-369, generating the construct pGBKT7-PrimPol 370-560 (Table 2.10). To test whether the PrimPol truncations fused to the GAL4 DNA binding domain result in auto-activation of the reporter genes, Y190 *S. cerevisae* were co-transformed with either pGBKT7-PrimPol 1-341 or 370-560 and empty pACT2, and cells containing both plasmids were tested for reporter gene transcription. Whilst cells transformed with pGBKT7:PrimPol 370-560 construct tested positive for activation of both reporter genes, cells transformed with pGBKT7:PrimPol 1-341 did not (Figure 7.10b). These cells did not test positive for β-galactosidase activity and were not able to grow on media containing 3-AT. Thus, the carboxyl-terminus of PrimPol containing the UL52 domain is responsible for auto-activation in the yeast-two hybrid assay.

7.5. Discussion

The experiments in this chapter reflect some of the initial attempts to discover the role of PrimPol in human cells. It was reasoned that by identifying proteins that interact with PrimPol, this would shed light on the cellular pathways in which this novel enzyme operates. The central approach adopted in this study was to stably express a *Strep*-tagged PrimPol in cultured human cells that could be affinity purified and then co-purifying proteins could be identified by mass-spectrometry analysis. However, the initial attempts documented here were not wholly successful. Far too many proteins were detected in the mass-spectrometry analysis indicating that a more stringent affinity purification process would be required in future, perhaps purifying PrimPol from a specific cellular compartment (nuclei or mitochondria) and using a two-step procedure taking advantage of both the Flag and *Strep*-Tag II epitopes. But even with increased stringency, it is not a certainty that PrimPol protein partners would be identified. It is not clear how 'active' PrimPol may be in an unperturbed cell, and so therefore may not be interacting with other proteins to perform its cellular role. It has become clear during the work performed in this thesis that PrimPol is involved in the tolerance of DNA damage, with a particular focus on DNA damage caused by UV-C irradiation, and so future experiments should affinity purify PrimPol from UV-C irradiated cells. Chapter 6 of this thesis showed that PrimPol becomes chromatin bound in 30-40 % of cells following irradiation with 20-30 J/m^2 UV-C, with binding occurring as quickly as 1 hour after irradiation; affinity purification of *Strep*-tagged PrimPol could be performed following a similar treatment. In addition, to remove the high level of over-expressed soluble PrimPol a chromatin preparation could be performed to enrich 'active' PrimPol, which can then be affinity purified. This could be done in one of two ways: following the enrichment of chromatin, DNA bound proteins could be released with a DNase (Benzonase) treatment and then affinity purification performed, or alternatively, a cross-linking approach could be applied similar to that used by Kannouche and colleagues in demonstrating Pol η's interaction with mono-ubiquitinated PCNA (Kannouche *et al*., 2004). Using one of these approaches on UV-C irradiated cells could identify PrimPol protein partners and shed light on the cellular processes this novel enzyme is involved in.

Despite the preliminary nature of the potential PrimPol protein partners obtained from the affinity purification mass-spectrometry experiment, some of these were re-produced in small-scale pull-downs. The single-stranded DNA binding proteins RPA and mtSSB both specifically co-purified with *Strep*-tagged PrimPol. Although little functional information can be gained from interaction with these two proteins given that single-stranded DNA binding proteins are involved in almost all DNA metabolic processes

(Oakley and Patrick, 2011). An important further experiment is to determine whether PrimPol's co-purification with these proteins is by protein-protein interaction or protein-DNA interaction. It could prove insightful to test the affect of RPA and mtSSB on PrimPol's catalytic activity *in vitro* as RPA was shown to be essential for Pol α-Prim-dependant primer synthesis and Okazaki fragment synthesis on DNA substrates (Matsumoto *et al*., 1990), and also, RPA along with PCNA was shown to promote error-free bypass of templated 8-oxo-G by a number of DNA polymerases (Maga *et al*., 2007). With PrimPol being implicated in the tolerance of UV-C induced DNA damage (Chapter 6), one very interesting hit is Pol δ interacting protein of 38 kDa, PDIP38 (also called PDIP2). As the name implies, PDIP38 associates with the replicative polymerase Pol δ in addition to PCNA (Liu *et al*., 2003), and is also a core component of mitochondrial nucleoids (Cheng *et al*., 2005). PDIP38 was recently shown to provide a novel link between replicative polymerases and TLS polymerases, interacting with and modulating the focal accumulation of Pol η (Tissier *et al*., 2010). It would be interesting to verify PrimPol's potential interaction with PDIP38, and test whether PDIP38 regulates the damage tolerance role of PrimPol in human cells.

Although no components of the mitochondrial replication machinery were identified in the preliminary mass-spectrometry analysis, a potential protein-protein interaction between PrimPol and the mtDNA replicase was identified using a far Western slot blot technique. Further experiments are required to verify this potential interaction, however it is interesting to speculate that interaction with Pol γ could allow PrimPol to become transiently associated with the mtDNA replisome and perform its yet to be fully characterised role in mtDNA metabolism. The yeast two-hybrid assay was also employed to identify PrimPol protein partners, however the carboxyl-terminus of PrimPol was found to cause auto-activation of the reporter genes. Further experiments could define the residues responsible for auto-activation within the carboxyl-terminus and the remainder of PrimPol could be used in a yeast-two hybrid assay, or alternatively, a mammalian two-hybrid could be attempted.

Protein-protein interactions play a fundamental role in cellular processes, and it is highly likely that PrimPol's interaction with other proteins is critical for its cellular role. Although the experiments in this chapter were not wholly successful, they lay the groundwork for further studies that could identify PrimPol protein partners.

Chapter 8

Conclusion

This thesis describes one of the first characterisations of the novel eukaryotic primase-polymerase PrimPol (also called CCDC111). At the outset of this work no reports were published regarding the function of this protein, only an *in silico* analysis by Iyer and colleagues (2005) that identified this protein as a putative member of the AEP superfamily. Subsequent unpublished work in the laboratories of Aidan Doherty and collaborator Luis Blanco (CBMSO, Madrid) confirmed PrimPol was an active AEP, and akin to its bacterial and archaeal counterparts, was a versatile nucleotidyl transferase, capable of *de novo* DNA-dependant DNA/RNA primase and DNA-dependant DNA/RNA polymerase activities *in vitro*. The general aim of this thesis was to elucidate the cellular role of this novel enzyme. A two-pronged approach was taken, characterising PrimPol in both human cells and the divergent eukaryotic model organism, and important human pathogen, the African trypanosome. Together these studies established a role for PrimPol in the cellular tolerance of DNA damage.

Whilst humans and most other eukaryotes encode a single PrimPol homologue, work detailed in Chapter 3 showed that trypanosomatids, a group of protozoan parasites responsible for a number of neglected human diseases, encode two PrimPol-like proteins. In the African trypanosome one of these PrimPol homologues, called TbPrimPol2, was demonstrated to be essential for cell survival. TbPrimPol2 localised to the nuclei of G2/M-phase cells and upon RNAi-mediated depletion, an irreversible cell cycle arrest occurred following the bulk of DNA synthesis. This cell cycle arrest coincided with the activation of the DNA damage response and was ultimately lethal. It was hypothesised that TbPrimPol2 was required for post-replication bypass of endogenously occurring DNA lesions, and consistent with this, Chapter 4 demonstrated that TbPrimPol2 (and its non essential paralogue TbPrimPol1) was capable of trans-lesion DNA synthesis *in vitro*. Both TbPrimPol1 and 2 could synthesise DNA opposite lesions that stall replicative DNA polymerases, such as the UV light-induced 6-4 photoproduct. Taken together, these data are consistent with a novel role for an AEP, in that is required for the completion of DNA replication by the post-replicative bypass of naturally occurring DNA damage. This also indicates that the mechanism of DNA replication in *T. brucei* has not only diverged with respect to the early steps of initiation, but also a downstream process.

The second half of this thesis focused on human PrimPol, with Chapter 5 detailing that unlike its essential *T. brucei* homologue, human PrimPol was present in both the nucleus and mitochondria of cells. Chapter 6 focused on the role of PrimPol in the nucleus and established that human PrimPol was also involved in DNA damage

tolerance. PrimPol re-localised in the nucleus following UV-C irradiation and assembled into numerous discrete foci tightly associated with chromatin. Preliminary data suggested these foci were largely restricted to S-phase cells and may represent replication factories stalled at UV photoproducts. RNAi-mediated depletion of PrimPol in normal human fibroblasts revealed increased chromatin-bound RPA and over-activation of the intra-S checkpoint following UV-C irradiation. This is indicative of increased replication fork stalling due to a defect in DNA damage tolerance. RNAi-mediated depletion of PrimPol in XP-V patient cells, which are deficient in the major UV photoproduct bypass polymerase Pol η, exacerbated the over-activation of the intra-S checkpoint following exposure to UV light. Although PrimPol depletion did not render normal cells sensitive to UV-C irradiation, it did sensitise XP-V cells, indicating PrimPol operates in a DNA damage tolerance pathway independent of Pol η, and that this pathway has increased importance in XP-V patient cells.

Although both human PrimPol and the *T. brucei* homologue TbPrimPol2 are implicated in DNA damage tolerance, their cellular role within this pathway may differ. TbPrimPol2 is an essential enzyme whilst higher eukaryotic PrimPol is not (Julie Bianchi, Laura Bailey, Aidan Doherty, unpublished). It is possible that TbPrimPol2 represents a trypanosomatid-specific PrimPol, perhaps required due to some of the novel features of trypanosomatid DNA metabolism, and that TbPrimPol1 is in fact performing a similar role to human PrimPol in the African trypanosome. In Chapter 3, a minor DNA damage response was observed in the nuclei of TbPrimPol1 depleted cells indicating a possible nuclear role, although this enzyme did not re-localise into foci following treatment with replication fork stalling agents (Lucy Glover and David Horn, unpublished), as observed with both TbPrimPol2 and human PrimPol. Although human PrimPol does not have an essential role in cell proliferation as demonstrated for TbPrimPol2, a growth defect was observed following RNAi of PrimPol in normal human fibroblasts, indicating human PrimPol does have a function in unperturbed cells. Although, studies of PrimPol$^{-/-}$ chicken cells indicated this defect was in replication fork progression as replication fork speeds were slowed in these cells (Julie Bianchi and Aidan Doherty, unpublished), whilst TbPrimPol2 depeleted cells show no defect in S-phase progression. Further work is required to elucidate the role of this novel primase-polymerase in both human cells and the African trypanosome. Regarding TbPrimPol2, it remains to be established what the endogenously occurring DNA modification is that requires TbPrimPol2 activity. A remaining question of human PrimPol is the *in vivo* activity of this enzyme, whether DNA/RNA priming is relevant to its cellular function or if its DNA/RNA polymerase activity that is required, or even both. *In vitro* experiments with recombinant human PrimPol can

begin to determine its preferences of substrate but additional cellular studies would also be required. Considering that African trypanosomes depleted of TbPrimPol2 can efficiently duplicate the bulk of their genome, this suggests TbPrimPol2 is not functioning as a primase, at least in the canonical sense of re-initiating DNA synthesis, as this would not be required following bulk DNA synthesis. Given human PrimPol is also present within the mitochondrion, future studies should explore the role of PrimPol within this essential organelle. It is highly likely that PrimPol would be playing a similar role, and deducing the mechanism of PrimPol-dependant DNA damage tolerance may prove easier in this small plasmid genome.

Despite the outstanding questions, the work presented in this thesis has established the involvement of the novel eukaryotic primase-polymerase PrimPol, in the universal mechanism of DNA damage tolerance, from one of the earliest diverging eukaryotic organisms to man.

References

Albertella, M. R., Green, C. M., Lehmann, A. R., and O'Connor, M. J. (2005). A role for polymerase η in the cellular tolerance to cisplatin-induced damage. Cancer Research, 65(21), 9799-9806.

Allan, C., Burel, J. M., Moore, J., Blackburn, C., Linkert, M., Loynton, S., Macdonald, D., Moore, W. J., Neves, C., Patterson, A., Porter, M., Tarkowska, A., Loranger, B., Avondo, J., Lagerstedt, I., Lianas, L., Leo, S., Hands, K., Hay, R. T., Patwardhan, A., Best, C., Kleywegt, G. J., Zanetti, G., Swedlow, J. R. (2012). OMERO: flexible, model-driven data management for experimental biology. Nature Methods, 9(3), 245-253.

Alsford, S., Kawahara, T., Glover, L., and Horn, D. (2005). Tagging a *T. brucei* RRNA locus improves stable transfection efficiency and circumvents inducible expression position effects. Molecular and Biochemical Parasitology, 144(2), 142–148.

Alsford, S., and Horn, D. (2008). Single-locus targeting constructs for reliable regulated RNAi and transgene expression in *Trypanosoma brucei*. Molecular and Biochemical Parasitology, 161(1), 76–79.

Alsford, S., Turner, D. J., Obado, S. O., Sanchez-Flores, A., Glover, L., Berriman, M., Hertz-Fowler, C., and Horn, D. (2011). High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome. Genome Research, 21(6), 915-924.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research, 25(17), 3389–3402.

Amerik, A. Y., and Hochstrasser, M. (2004). Mechanism and function of deubiquitinating enzymes. Biochimica et Biophysica Acta, 1695(1-3), 189-207.

Andersen, P. L., Xu, F., Ziola, B., McGregor, W. G., and Xiao, W. (2011). Sequential assembly of translesion DNA polymerases at UV-induced DNA damage sites. Molecular Biology of the Cell, 22(13), 2373-83.

Araujo, S. J., Tirode, F., Coin, F., Pospiech, H., Syvaoja, J. E., Stucki, M., Hubscher, U., Egly, J. M. and Wood, R. D. (2000). Nucleotide excision repair of DNA with recombinant human proteins: definition of the minimal set of factors, active forms of TFIIH, and modulation by CAK. Genes and Development, 14, 349-359.

Aravind, L., Leipe, D. D., and Koonin, E. V. (1998). Toprim--a conserved catalytic domain in type IA and II topoisomerases, DnaG-type primases, OLD family nucleases and RecR proteins. Nucleic Acids Research, 26(18), 4205-13.

Aravind, L. and Koonin, E.V. (2001). Prokaryotic homologs of the eukaryotic DNA-end-binding protein Ku, novel domains in the Ku protein and prediction of a prokaryotic double-strand break repair system. Genome Research. 11(8), 1365-74.

Arezi, B., and Kuchta, R. D. (2000). Eukaryotic DNA primase. Trends in Biochemical Sciences, 25(11), 572–576.

Arlett, C. F., Harcourt, S. A., and Broughton, B. C. (1975). The influence of caffeine on cell survival in excision-proficient and excision-deficient xeroderma pigmentosum and normal human cell strains following ultraviolet-light irradiation. Mutation Research, 33, 341-346.

Aslett, M., Aurrecoechea, C., Berriman, M., Brestelli, J., Brunk, B. P., Carrington, M., Depledge, D. P., Fischer, S., Gajria, B., Gao, X., Gardner, M. J., Gingle, A., Grant, G., Harb, O. S., Heiges, M., Hertz-Fowler, C., Houston, R., Innamorato, F., Iodice, J., Kissinger, J. C., Kraemer, E., Li, W., Logan, F. J., Miller, J. A., Mitra, S., Myler, P. J., Nayak, V., Pennington, C., Phan, I., Pinney, D. F., Ramasamy, G., Rogers, M. B., Roos, D. S., Ross, C., Sivam, D.,

Smith, D. F., Srinivasamoorthy, G., Stoeckert, C. J. Jr, Subramanian, S., Thibodeau, R., Tivey, A., Treatman, C., Velarde, G., and Wang, H. (2010). TriTrypDB: a functional genomic resource for the Trypanosomatidae. Nucleic Acids Research, 38, D457–62.

Augustin, M. A., Huber, R. and Kaiser, J. T. (2001). Crystal structure of a DNA-dependent RNA polymerase (DNA primase). Nature Structural and Molecular Biology, 8(1), 57-61.

Avkin S., and Livneh Z., (2002). Efficiency, specificity and DNA polymerase-dependence of translesion replication across the oxidative DNA lesion 8-oxoguanine in human cells. Mutation Research, 510, 81–90.

Avkin, S., Goldsmith, M., Velasco-Miguel, S., Geacintov, N., Friedberg, E. C., and Livneh, Z. (2004). Quantitative analysis of translesion DNA synthesis across a benzo[a]pyrene-guanine adduct in mammalian cells: the role of DNA polymerase κ. The Journal of Biological Chemistry, 279(51), 53298-305.

Avkin, S., Sevilya, Z., Toube, L., Geacintov, N., Chaney, S. G., Oren, M., and Livneh, Z. (2006). P53 and p21 regulate error-prone DNA repair to yield a lower mutation load. Molecular Cell, 22, 407–413.

Bambara, R. A., Fay, P. J., and Mallaber, L. M. (1995). Methods of analyzing processivity. Methods in enzymology, 262, 270-80.

Baranovskiy, A. G., Lada, A. G., Siebler, H. M., Zhang, Y., Pavlov, Y. I., and Tahirov, T. H. (2012). DNA polymerase δ and ζ switch by sharing accessory subunits of DNA polymerase δ. Journal of Biological Chemistry, 287(21), 17281-7.

Barrett, M. P., Burchmore, R. J., Stich, A., Lazzari, J. O., Frasch, A. C., Cazzulo, J. J., and Krishna, S. (2003). The trypanosomiases. Lancet, 362(9394), 1469-80.

Batista, L. F. Z., Kaina, B., Meneghini, R., and Menck, C. F. M. (2009). How DNA lesions are turned into powerful killing structures: Insights from UV-induced apoptosis. Mutation Research, 681(2-3), 197–208.

Bailly, V., Lauder, S., Prakash, S. and Prakash, L. (1997). Yeast DNA repair proteins Rad6 and Rad18 form a heterodimer that has ubiquitin conjugating, DNA binding, and ATP hydrolytic activities. The Journal of Biological Chemistry, 272, 23360-23365.

Beard, W. A., and Wilson, S. H. (2001). DNA lesion bypass polymerases open up. Structure, 9(9), 759-64.

Beard, W. A., Shock, D. D., Vande Berg, B. J., and Wilson, S. H. (2002). Efficiency of correct nucleotide insertion governs DNA polymerase fidelity. The Journal of Biological Chemistry, 277(49), 47393-8.

Beattie, T. R., and Bell, S. D. (2012). Coordination of multiple enzyme activities by a single PCNA in archaeal Okazaki fragment maturation. EMBO Journal, 31(6), 1556-67.

Bebenek, K., Tissier, A., Frank, E. G., McDonald, J. P., Prasad, R., Wilson, S. H., Woodgate, R., and Kunkel, T. A. (2001). 5'-Deoxyribose phosphate lyase activity of human DNA polymerase iota in vitro. Science, 291(5511), 2156-9.

Beck, K., and Lipps, G. (2007). Properties of an unusual DNA primase from an archaeal plasmid. Nucleic Acids Research, 35(17), 5635-45.

Beese, L. S., and Steitz, T. A. (1991). Structural basis for the 3'-5' exonuclease activity of *Escherichia coli* DNA polymerase I: a two metal ion mechanism. The EMBO Journal, 10(1), 25-33.

Bell, S. P., and Dutta, A. (2002). DNA replication in eukaryotic cells. Annual Review of Biochemistry, 71, 333-74.

Bell, J. S., Harvey, T. I., Sims, A. M., and McCulloch, R. (2004). Characterization of components of the mismatch repair machinery in Trypanosoma brucei. Molecular Microbiology, 51(1), 159-73.

Bemark, M., Khamlichi, A. A., Davies, S. L., and Neuberger, M. S. (2000). Disruption of mouse polymerase ζ (Rev3) leads to embryonic lethality and impairs blastocyst development in vitro. Current Biology, 10(19), 1213-6.

Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., Sayers, E. W. (2009). GenBank. Nucleic Acids Research, 37, D26-31.

Berquist, B. R., and Wilson, D. M. (2012). Pathways for repairing and tolerating the spectrum of oxidative DNA lesions. Cancer Letters, 327(1-2), 61-72.

Bermejo, R., Lai, M. S., and Foiani, M. (2012). Preventing replication stress to maintain genome stability: resolving conflicts between replication and transcription. Molecular Cell, 45(6), 710-8.

Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D. C., Lennard, N. J., Caler, E., Hamlin, N. E., Haas, B., Böhme, U., Hannick, L., Aslett, M. A., Shallom, J., Marcello, L., Hou, L., Wickstead, B., Alsmark, U. C., Arrowsmith, C., Atkin, R. J., Barron, A. J., Bringaud, F., Brooks, K., Carrington, M., Cherevach, I., Chillingworth, T. J., Churcher, C., Clark, L. N., Corton, C. H., Cronin, A., Davies, R. M., Doggett, J. , Djikeng, A., Feldblyum, T., Field, M. C., Fraser, A., Goodhead, I., Hance, Z., Harper, D., Harris, B. R., Hauser, H., Hostetler, J., Ivens, A., Jagels, K., Johnson, D., Johnson, J., Jones, K., Kerhornou, A. X., Koo, H., Larke, N., Landfear, S., Larkin, C., Leech, V., Line, A., Lord, A., Macleod, A., Mooney, P. J., Moule, S., Martin, D. M., Morgan, G. W., Mungall, K., Norbertczak, H., Ormond, D., Pai, G., Peacock, C. S., Peterson, J., Quail, M. A., Rabbinowitsch, E., Rajandream, M. A., Reitter, C., Salzberg, S. L., Sanders, M., Schobel, S., Sharp, S., Simmonds, M., Simpson, A. J., Tallon, L., Turner, C. M., Tait, A., Tivey, A. R., Van Aken, S., Walker, D., Wanless, D., Wang, S., White, B., White, O., Whitehead, S., Woodward, J., Wortman, J., Adams, M. D., Embley, T. M., Gull, K., Ullu, E., Barry, J. D., Fairlamb, A. H., Opperdoes, F., Barrell, B. G., Donelson, J. E., Hall, N., Fraser, C. M., Melville, S. E., El-Sayed, N. M. (2005). The genome of the African trypanosome Trypanosoma brucei. Science, 309(5733), 416-22.

Bétous, R., Rey, L., Wang, G., Pillaire, M. J., Puget, N., Selves, J., Biard, D. S., Shin-ya, K., Vasquez, K. M., Cazaux, C., and Hoffmann, J. S. (2009). Role of TLS DNA polymerases η and κ in processing naturally occurring structured DNA in human cells. Molecular Carcinogenesis, 48(4), 369-78.

Bi, X., Slater, D. M., Ohmori, H., and Vaziri, C. (2005). DNA polymerase κ is specifically required for recovery from the benzo[a]-pyrene-dihydrodiol epoxide (BPDE)-induced S-phase checkpoint. The Journal of Biological Chemistry, 280, 22343–22355.

Bienko, M., Green, C. M., Crosetto, N., Rudolf, F., Zapart, G., Coull, B., Kannouche, P., Wider, G., Peter, M., Lehmann, A. R., Hofmann, K., and Dikic, I. (2005). Ubiquitin-binding domains in Y-family polymerases regulate translesion synthesis. Science, 310(5755), 1821-4.

Bienko, M., Green, C. M., Sabbioneda, S., Crosetto, N., Matic, I., Hibbert, R. G., Begovic, T., Niimi, A., Mann, M., Lehmann, A. R., and Dikic, I. (2010). Regulation of translesion synthesis DNA polymerase eta by monoubiquitination. Molecular Cell, 37(3), 396-407.

Biertümpfel, C., Zhao, Y., Kondo, Y., Ramón-Maiques, S., Gregory, M., Lee, J. Y., Masutani, C., Lehmann, A. R., Hanaoka, F., and Yang, W. (2010). Structure and mechanism of human DNA polymerase η. Nature, 465(7301), 1044-8.

Biswas, N., and Weller, S. K. (1999). A mutation in the C-terminal putative Zn^{2+} finger motif of UL52 severely affects the biochemical activities of the HSV-1 helicase-primase subcomplex. Journal of Biological Chemistry, 274(12), 8068-76.

Bocquier, A. A., Liu, L., Cann, I. K., Komori, K., Kohda, D., and Ishino, Y. (2001). Archaeal primase: bridging the gap between RNA and DNA polymerases. Current Biology, 11(6), 452-6.

Bollum, F. J., and Potter, V. R. (1958). Incorporation of thymidine into deoxyribonucleic acid by enzymes from rat tissues. The Journal of Biological Chemistry, 233(2), 478-82.

Bomgarden, R. D., Lupardus, P. J., Soni, D. V., Yee, M.-C., Ford, J. M., and Cimprich, K. A. (2006). Opposing effects of the UV lesion repair protein XPA and UV bypass polymerase eta on ATR checkpoint signaling. The EMBO Journal, 25(11), 2605–2614.

Boudsocq, F., Kokoska, R. J., Plosky, B. S., Vaisman, A., Ling, H., Kunkel, T. A., Yang, W., and Woodgate, R. (2004). Investigating the role of the little finger domain of Y-family DNA polymerases in low fidelity synthesis and translesion replication. The Journal of Biological Chemistry, 279(31), 32932-40.

Boyce, R.P., Howard-Flanders, P. (1964). Release of ultraviolet light-induced thymine dimers from DNA in E. coli K-12, Proceedings of the National Academy of Sciences, 51, 293–300.

Brautigam, C. A., and Steitz, T. A. (1998). Structural and functional insights provided by crystal structures of DNA polymerases and their substrate complexes. Current Opinion in Structural Biology, 8(1), 54-63.

Breslauer, K. J., Frank, R., Blöcker, H., and Marky, L. A. (1986). Predicting DNA duplex stability from the base sequence. Proceedings of the National Academy of Sciences, 83(11), 3746-3750.

Bridges, B. A. (2005). Error-prone DNA repair and translesion synthesis: focus on the replication fork. DNA Repair, 4(5), 618-9, 634.

Brissett, N. C., Pitcher, R. S., Juarez, R., Picher, A. J., Green, A. J., Dafforn, T. R., Fox, G. C., Blanco, L., and Doherty, A. J. (2007). Structure of a NHEJ polymerase-mediated DNA synaptic complex. Science, 318(5849), 456-9.

Brissett, N. C., and Doherty, A. J. (2009). Repairing DNA double-strand breaks by the prokaryotic non-homologous end-joining pathway. Biochemical Society Transactions, 37(3), 539–545.

Brissett, N. C., Martin, M. J., Pitcher, R. S., Bianchi, J., Juarez, R., Green, A. J., Fox, G. C., Blanco, L., and Doherty, A. J. (2011). Structure of a Preternary Complex Involving a Prokaryotic NHEJ DNA Polymerase. Molecular Cell, 41(2), 221–231.

Brun, R., Blum, J., Chappuis, F., and Burri, C. (2010). Human African trypanosomiasis. Lancet, 375(9709), 148-59.

Brutlag, D., Schekman, R., and Kornberg, A. (1971). A possible role for RNA polymerase in the initiation of M13 DNA synthesis. Proceedings of the National Academy of Sciences, 68(11), 2826-9.

Bullock, S. K., Kaufmann, W. K., and Cordeiro-Stone, M. (2001). Enhanced S phase delay and inhibition of replication of an undamaged shuttle vector in UVC-irradiated xeroderma pigmentosum variant. Carcinogenesis, 22(2), 233-241.

Burgers, P. M. (1991). Saccharomyces cerevisiae replication factor C. II. Formation and activity of complexes with the proliferating cell nuclear antigen and with DNA polymerases δ and ε. Journal of Biological Chemistry, 266(33), 22698-706.

Burgers, P. M. J. (2008). Polymerase Dynamics at the Eukaryotic DNA Replication Fork. The Journal of Biological Chemistry, 284(7), 4041–4045.

Burton, P., McBride, D. J., Wilkes, J. M., Barry, J. D., and McCulloch, R. (2007). Ku heterodimer-independent end joining in Trypanosoma brucei cell extracts relies upon sequence microhomology. Eukaryotic Cell, 6(10), 1773-1781.

Cadet, J., Douki, T., Gasparutto, D., and Ravanat, J. L. (2003). Oxidative damage to DNA: formation, measurement and biochemical features. Mutation Research, 531(1-2), 5-23.

Cairns, J. (1963). The bacterial chromosome and its manner of replication as seen by autoradiography. Journal of Molecular Biology, 6, 208-13.

Callegari, A. J., and Kelly, T. J. (2006). UV irradiation induces a postreplication DNA damage checkpoint. Proceedings of the National Academy of Sciences. 103(43), 15877-15882.

Callegari, A. J., Clark, E., Pneuman, A., and Kelly, T. J. (2010). Postreplication gaps at UV lesions are signals for checkpoint activation. Proceedings of the National Academy of Sciences, 107(18), 8219–8224.

Capasso, H., Palermo, C., Wan, S., Rao, H., John, U. P., O'Connell, M. J., and Walworth, N. C. (2002). Phosphorylation activates Chk1 and is required for checkpoint-mediated cell cycle arrest. Journal of Cell Science, 115(23), 4555-4564.

Chang, D. J., Lupardus, P. J., and Cimprich, K. A. (2006). Monoubiquitination of proliferating cell nuclear antigen induced by stalled replication requires uncoupling of DNA polymerase and mini-chromosome maintenance helicase activities. The Journal of Biological Chemistry, 281(43), 32081-8.

Chang, D. J., and Cimprich, K. A. (2009). DNA damage tolerance: when it's OK to make mistakes. Nature Chemical Biology, 5(2), 82–90.

Chen, Y., Carrington-Lawrence, S. D., Bai, P., and Weller, S. K. (2005). Mutations in the putative zinc-binding motif of UL52 demonstrate a complex interdependence between the UL5 and UL52 subunits of the human herpes simplex virus type 1 helicase/primase complex. Journal of Virology, 79(14), 9088-96.

Cheng, X., Kanki, T., Fukuoh, A., Ohgaki, K., Takeya, R., Aoki, Y., Hamasaki, N., and Kang, D. (2005). PDIP38 associates with proteins constituting the mitochondrial DNA nucleoid. The Journal of Biochemistry, 138(6), 673-678.

Chilkova, O., Stenlund, P., Isoz, I., Stith, C. M., Grabowski, P., Lundström, E. B., Burgers, P. M., and Johansson, E. (2007). The eukaryotic leading and lagging strand DNA polymerases are loaded onto primer-ends via separate mechanisms but have comparable processivity in the presence of PCNA. Nucleic Acids Research, 35(19), 6588-97.

Chiu, R. K., Brun, J., Ramaekers, C., Theys, J., Weng, L., Lambin, P., Gray, D. A., and Wouters, B. G. (2006). Lysine 63-polyubiquitination guards against translesion synthesis-induced mutations. PLOS Genetics, 2(7), e116.

Chung, C. T., Niemela, S. L., and Miller, R. H. (1989). One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. Proceedings of the National Academy of Sciences of the United States of America, 86(7), 2172–2175.

Cleaver, J. E., Afzal, V., Feeney, L., McDowell, M., Sadinski, W., Volpe, J. P., Busch, D. B., Coleman, D. M., Ziffer, D. W., Yu, Y., Nagasawa, H., Little, J. B. (1999). Increased ultraviolet sensitivity and chromosomal instability related to P53 function in the xeroderma pigmentosum variant. Cancer Research, 59(5), 1102-1108.

Cline, S. D. (2012). Mitochondrial DNA damage and its consequences for mitochondrial gene expression. Biochimica et Biophysica Acta, 1819(9-10), 979-991.

Conway, C., McCulloch, R., Ginger, M. L., Robinson, N. P., Browitt, A., and Barry, J. D. (2002a). Ku is important for telomere maintenance, but not for differential expression of telomeric VSG genes, in African trypanosomes. Journal of Biological Chemistry, 277(24), 21269-77.

Conway, C., Proudfoot, C., Burton, P., Barry, J. D., and McCulloch, R. (2002b). Two pathways of homologous recombination in *Trypanosoma brucei*. Molecular Microbiology, 45(6), 1687-700.

Copeland, W. C., and Tan, X. (1995). Active site mapping of the catalytic mouse primase subunit by alanine scanning mutagenesis. Journal of Biological Chemistry, 270(8), 3905-13.

Cordeiro-Stone, M., Frank, A., Bryant, M., Oguejiofor, I., Hatch, S. B., McDaniel, L. D., and Kaufmann, W. K. (2002). DNA damage responses protect xeroderma pigmentosum variant from UVC-induced clastogenesis. Carcinogenesis, 23(6), 959-965.

Corn, J. E., Pelton, J. G., and Berger, J. M. (2006). Identification of a DNA primase template tracking site redefines the geometry of primer synthesis. Nature Structural and Molecular Biology, 15(2), 163-9.

Courcelle, J., Donaldson, J. R., Chow, K. H., and Courcelle, C. T. (2003). DNA damage-induced replication fork regression and processing in Escherichia coli. Science, 299(5609), 1064-7.

Cox, B. S, and Parry, J. M. (1968). The isolation, genetics and survival characteristics of ultraviolet light-sensitive mutants in yeast. Mutation Research, 6(1), 37-55.

Cuvier, O., Stanojcic, S., Lemaitre, J. M. and Mechali, M. (2008). A topoisomerase II-dependent mechanism for resetting replicons at the S-M-phase transition. Genes and Development, 22(7), 860-5.

Daigaku, Y., Davies, A. A. and Ulrich, H. D. (2010). Ubiquitin- dependent DNA damage bypass is separable from genome replication. Nature, 465(7300), 951-5.

Daigaku, Y. (2012). Roadworks of DNA Damage Bypass during and after Replication. Genes and Environment, 34(2), 77-88.

Dandekar, T., Snel, B., Huynen, M., and Bork, P. (1998). Conservation of gene order: a fingerprint of proteins that physically interact. Trends in Biochemical Sciences, 23, 324–28.

Dang, H. Q., and Li, Z. (2011). The Cdc45•Mcm2-7•GINS protein complex in trypanosomes regulates DNA replication and interacts with two Orc1-like proteins in the origin recognition complex. Journal of Biological Chemistry, 286(37), 32424-35.

Daniels, J.-P., Gull, K., and Wickstead, B. (2010). Cell biology of the trypanosome genome. Microbiology and Molecular Biology Reviews, 74(4), 552–569.

Davey, S. K., and Faust, E. A. (1990). Murine DNA polymerase alpha-primase initiates RNA-primed DNA synthesis preferentially upstream of a 3′-CC(C/A)-5′ motif. Journal of Biological Chemistry, 265(7), 3611-4.

Davies, J. F. 2nd, Almassy, R. J., Hostomska, Z., Ferre, R. A., and Hostomsky, Z. (1994). 2.3 A crystal structure of the catalytic domain of DNA polymerase beta. Cell, 76(6), 1123-33.

Davies, A. A., Huttner, D., Daigaku, Y., Chen, S., and Ulrich, H. D. (2008). Activation of ubiquitin-dependent DNA damage bypass is mediated by replication protein a. Molecular Cell, 29(5), 625-36.

De Falco, M., Fusco, A., De Felice, M., Rossi, M., and Pisani, F. M. (2004). The DNA primase of Sulfolobus solfataricus is activated by substrates containing a thymine-rich bubble and has a 30-terminal nucleotidyl-transferase activity. Nucleic Acids Research, 32, 5223–5230.

de Feraudy, S., Limoli, C. L., Giedzinski, E., Karentz, D., Marti, T. M., Feeney, L., and Cleaver, J. E. (2007). Pol η is required for DNA replication during nucleotide deprivation by hydroxyurea. Oncogene, 26(39), 5713–5721.

de Moura, M. B., Schamber-Reis, B. L., Passos Silva, D. G., Rajão, M. A., Macedo, A. M., Franco, G. R., Pena, S. D., Teixeira, S. M., and Machado, C. R. (2009). Cloning and characterization of DNA polymerase η from *Trypanosoma cruzi*: roles for translesion bypass of oxidative damage. Environmental and Molecular Mutagenesis, 50(5), 375-86.

de Padula, M., Slezak, G., Auffret van Der Kemp, P., and Boiteux, S. (2004). The post-replication repair RAD18 and RAD6 genes are involved in the prevention of spontaneous mutations caused by 7,8-dihydro-8-oxoguanine in *Saccharomyces cerevisiae*. Nucleic Acids Research, 32(17), 5003-5010.

Della, M., Palmbos, P. L., Tseng, H. M., Tonkin, L. M., Daley, J. M., Topper, L. M., Pitcher, R. S., Tomkinson, A. E., Wilson, T. E., and Doherty, A. J. (2004). Mycobacterial Ku and ligase proteins constitute a two-component NHEJ repair machine. Science, 306(5696), 683-5.

Dennis, G Jr., Sherman, B. T., Hosack, D. A., Yang, J., Gao, W., Lane, H. C., and Lempicki, R. A. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biology, 4(5), 3.

Despras, E., Daboussi, F., Hyrien, O., Marheineke, K., and Kannouche, P. L. (2010). ATR/Chk1 pathway is essential for resumption of DNA synthesis and cell survival in UV-irradiated XP variant cells. Human Molecular Genetics, 19(9), 1690–1701.

Despras, E., Delrieu, N., Garandeau, C., Ahmed-Seghir, S., and Kannouche, P. L. (2012). Regulation of the specialized DNA polymerase eta: Revisiting the biological relevance of its PCNA- and ubiquitin-binding motifs. Environmental and Molecular Mutagenesis, 53(9), 752-65.

di Caprio, L., and Cox, B. S. (1981). DNA synthesis in UV-irradiated yeast. Mutation Research, 82(1), 69-85.

Di Re, M., Sembongi, H., He, J., Reyes, A., Yasukawa, T., Martinsson, P., Bailey, L. J., Goffart, S., Boyd-Kirkup, J. D., Wong, T. S., Fersht, A. R., Spelbrink, J. N., and Holt, I. J. (2009). The accessory subunit of mitochondrial DNA polymerase γ determines the DNA content of mitochondrial nucleoids in human cultured cells. Nucleic Acids Research, 37(17), 5701-5713.

Diamant, N., Hendel, A., Vered, I., Carell, T., Reissner, T., de Wind, N., Geacinov, N., and Livneh, Z. (2012). DNA damage bypass operates in the S and G2 phases of the cell cycle and exhibits differential mutagenicity. Nucleic Acids Research, 40, 170–180.

DiNardo, S., Voelkel, K. and Sternglanz, R. (1984). DNA topoisomerase II mutant of Saccharomyces cerevisiae: topoisomerase II is required for segregation of daughter molecules at the termination of DNA replication. Proceedings of the National Academy of Sciences, 81(9), 2616-20.

Doherty, A. J., Ashford, S. R., Brannigan, J. A., and Wigley, D. B. (1995). A superior host strain for the over-expression of cloned genes using the T7 promoter based vectors. Nucleic Acids Research, 23(11), 2074–2075.

Doherty, A. J., Jackson, S. P., and Weller, G. R. (2001). Identification of bacterial homologues of the Ku DNA repair proteins. FEBS Letters. 500, 186–88.

Dulbecco, R. (1949). Reactivation of ultra-violet-inactivated bacteriophage by visible light. Nature, 162,949–950.

Dumstorf, C. A., Clark, A. B., Lin, Q., Kissling, G. E., Yuan, T., Kucherlapati, R., McGregor, W. G., and Kunkel, T. A. (2006). Participation of mouse DNA polymerase ι in strand-biased mutagenic bypass of UV photoproducts and suppression of skin cancer. Proceedings of the National Academy of Sciences, 103 18083–18088.

Durfee, T., Becherer, K., Chen, P. L., Yeh, S. H., Yang, Y., Kilburn, A. E., Lee, W. H., and Elledge, S. J. (1993). The retinoblastoma protein associates with the protein phosphatase type 1 catalytic subunit. Genes and Development, 7(4), 555-569.

Edgar, R. C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics, 5, 113.

Edmunds, C. E., Simpson, L. J., and Sale, J. E. (2008) PCNA ubiquitination and REV1 define temporally distinct mechanisms for controlling translesion synthesis in the avian cell line DT40. Molecular Cell, 30(4), 519-529.

Eid, J., and Sollnerwebb, B. (1991). Stable integrative transformation of *Trypanosoma brucei* that occurs exclusively by homologous recombination. Proceedings of the National Academy of Sciences, 88, 2118–2121.

El-Sayed, N. M., Hegde, P., Quackenbush, J., Melville, S. E., and Donelson, J. E. (2000). The African trypanosome genome. International Journal for Parasitology, 30(4), 329-45.

El-Sayed, N. M., Myler, P. J., Bartholomeu, D. C., Nilsson, D., Aggarwal, G., Tran, A. N., Ghedin, E., Worthey, E. A., Delcher, A. L., Blandin, G., Westenberger, S. J., Caler, E., Cerqueira, G. C., Branche, C., Haas, B., Anupama, A., Arner, E., Aslund, L., Attipoe, P., Bontempi, E., Bringaud, F., Burton, P., Cadag, E., Campbell, D. A., Carrington, M., Crabtree, J., Darban, H., da Silveira, J. F., de Jong, P., Edwards, K., Englund, P. T., Fazelina, G., Feldblyum, T., Ferella, M., Frasch, A. C., Gull, K., Horn, D., Hou, L., Huang, Y., Kindlund, E., Klingbeil, M., Kluge, S., Koo, H., Lacerda, D., Levin, M. J., Lorenzi, H., Louie, T., Machado, C. R., McCulloch, R., McKenna, A., Mizuno, Y., Mottram, J. C., Nelson, S., Ochaya, S., Osoegawa, K., Pai, G., Parsons, M., Pentony, M., Pettersson, U., Pop, M., Ramirez, J. L., Rinta, J., Robertson, L., Salzberg, S. L., Sanchez, D. O., Seyler, A., Sharma, R., Shetty, J., Simpson, A. J., Sisk, E., Tammi, M. T., Tarleton, R., Teixeira, S., Van Aken, S., Vogt, C., Ward, P. N., Wickstead, B., Wortman, J., White, O., Fraser, C. M., Stuart, K. D., and Andersson, B. (2005). The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. Science, 309(5733), 409-15.

Elvers, I., Johansson, F., Groth, P., Erixon, K., and Helleday, T. (2011). UV stalled replication forks restart by re-priming in human fibroblasts. Nucleic Acids Research, 39(16), 7049-57.

Errico, A., and Costanzo, V. (2010). Differences in the DNA replication of unicellular eukaryotes and metazoans: known unknowns. EMBO Reports, 11(4), 270-8.

Esposito, G., Godindagger, I., Klein, U., Yaspo, M. L., Cumano, A., and Rajewsky, K. (2000). Disruption of the Rev3l-encoded catalytic subunit of polymerase ζ in mice results in early embryonic lethality. Current Biology, 10(19), 1221-4.

Fachinetti, D., Bermejo, R., Cocito, A., Minardi, S., Katou, Y., Kanoh, Y., Shirahige, K., Azvolinsky, A., Zakian, V. A., and Foiani, M. (2010). Replication termination at eukaryotic chromosomes is mediated by Top2 and occurs at genomic loci containing pausing elements. Molecular Cell, 39(4), 595-605.

Fairman, M. P., and Stillman, B. (1988). Cellular factors required for multiple stages of SV40 DNA replication in vitro. The EMBO Journal, (4), 1211-8.

Fields, S., and Song, O. (1989). A novel genetic system to detect protein-protein interactions. Nature, 340(6230), 245-246.

Fields-Berry, S. C. and DePamphilis, M. L. (1989). Sequences that promote formation of catenated intertwines during termination of DNA replication. Nucleic Acids Research, 17, 3261–3273.

Foiani, M., Santocanale, C., Plevani, P., and Lucchini G. (1989). A single essential gene, PRI2, encodes the large subunit of DNA primase in *Saccharomyces cerevisiae*. Molecular and Cellular Biology, 9(7), 3081-3087.

Foiani, M., Lucchini, G., and Plevani, P. (1997). The DNA polymerase alpha-primase complex couples DNA replication, cell-cycle progression and DNA-damage response. Trends in Biochemical Sciences, 22(11), 424–427.

Frick, D. N., Kumar, S., and Richardson, C. C. (1999). Interaction of ribonucleoside triphosphates with the gene 4 primase of bacteriophage T7. Journal of Biological Chemistry, 274(50), 35899-907.

Frick, D. N., and Richardson, C. C. (2001). DNA primases. Annual Review of Biochemistry, 70, 39-80.

Fu, D., Calvo, J. A., and Samson, L. D. (2012). Balancing repair and tolerance of DNA damage caused by alkylating agents. Nature Reviews Cancer, 12(2), 104-120.

Fusté, J. M., Wanrooij, S., Jemt, E., Granycome, C. E., Cluett, T. J., Shi, Y., Atanassova, N., Holt, I. J., Gustafsson, C. M., and Falkenberg, M. (2010). Mitochondrial RNA polymerase is needed for activation of the origin of light-strand DNA replication. Molecular cell, 37(1), 67–78.

Gan, G. N., Wittschieben, J. P., Wittschieben, B. Ø., and Wood, R. D. (2008). DNA polymerase zeta (pol ζ) in higher eukaryotes. Cell Research, 18(1), 174-83.

Garg, P., and Burgers, P. M. (2005a). DNA polymerases that propagate the eukaryotic DNA replication fork. Critical Reviews in Biochemistry and Molecular Biology, 40(2), 115-28.

Garg, P., and Burgers, P. M. (2005b). Ubiquitinated proliferating cell nuclear antigen activates translesion DNA polymerases η and REV1. Proceedings of the National Academy of Sciences, 102 (2005), 18361–18366.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., and Bairoch, A. (2003). ExPASy: The proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Research, 31(13), 3784-3788.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. The Proteomics Protocols Handbook, Humana Press. 571-607.

Gerlach, V. L., Aravind, L., Gotway, G., Schultz, R. A., Koonin, E. V., and Friedberg, E. C. (1999). Human and mouse homologs of Escherichia coli DinB (DNA polymerase IV), members of the UmuC/DinB superfamily. Proceedings of the National Academy of Sciences, 96(21), 11922-7.

Gibbs, P. E., McGregor, W. G., Maher, V. M., Nisson, P., and Lawrence, C. W. (1998). A human homolog of the Saccharomyces cerevisiae REV3 gene, which encodes the catalytic subunit of DNA polymerase ζ. Proceedings of the National Academy of Sciences, 95(12), 6876-80.

Gibbs, P. E., Wang, X. D., Li, Z., McManus, T. P., McGregor, W. G., Lawrence, C. W., and Maher, V. M. (2000). The function of the human homolog of *Saccharomyces cerevisiae* REV1 is required for mutagenesis induced by ultraviolet light. Proceedings of the National Academy of Sciences, 97(8), 4186-91.

Glover, L., McCulloch, R., and Horn, D. (2008). Sequence homology and microhomology dominate chromosomal double-strand break repair in African trypanosomes. Nucleic Acids Research, 36(8), 2608–2618.

Glover, L., and Horn, D. (2012). Trypanosomal histone γH2A and the DNA damage response. Molecular and Biochemical Parasitology, 183(1), 78–83.

Godoy, P. D., Nogueira-Junior, L. A., Paes, L. S., Cornejo, A., Martins, R. M., Silber, A. M., Schenkman, S., and Elias, M. C. (2009). Trypanosome prereplication machinery contains a single functional orc1/cdc6 protein, which is typical of archaea. Eukaryotic Cell, 8(10), 1592-603.

Godoy, V. G., Jarosz, D. F., Walker, F. L., Simmons, L. A., and Walker, G. C. (2006). Y-family DNA polymerases respond to DNA damage-independent inhibition of replication fork progression. The EMBO Journal, 25, 868–879.

Göhler, T., Sabbioneda, S., Green, C. M., and Lehmann, A. R. (2011). ATR-mediated phosphorylation of DNA polymerase η is needed for efficient recovery from UV damage. Journal of Cell Biology, 192(2), 219-27.

Goodman, M. F. (1997). Hydrogen bonding revisited: geometric selection as a principal determinant of DNA replication fidelity. Proceedings of the National Academy of Sciences, 94(20), 10493-5.

Goulian, M., and Kornberg, A. (1967). Enzymatic synthesis of DNA. 23. Synthesis of circular replicative form of phage phi-X174 DNA. Proceedings of the National Academy of Sciences, 58(4), 1723-30.

Goulian, M., Lucas, Z. J., and Kornberg, A. (1968). Enzymatic synthesis of deoxyribonucleic acid. XXV. Purification and properties of deoxyribonucleic acid polymerase induced by infection with phage T4. Journal of Biological Chemistry, 243(3), 627-38.

Graziewicz, M. A., Day, B. J., and Copeland, W. C. (2002). The mitochondrial DNA polymerase as a target of oxidative damage. Nucleic Acids Research, 30(13), 2817-2824

Graziewicz, M. A., Bienstock, R. J., and Copeland, W. C. (2007) The DNA polymerase g Y955C disease variant associated with PEO and parkinsonism mediates the incorporation and translesion synthesis opposite 7,8-dihydro-8-oxo-20-deoxyguanosine. Human Molecular Genetics, 16(22), 2729-2739.

Grosse, F., and Krauss, G. (1985). The primase activity of DNA polymerase α from calf thymus. Journal of Biological Chemistry, 260(3), 1881-8.

Gueranger, Q., Stary, A., Aoufouchi, S., Faili, A., Sarasin, A., Reynaud, C. A., and Weill, J. C. (2008). Role of DNA polymerases η, ι and ζ in UV resistance and UV-induced mutagenesis in a human cell line. DNA Repair, 7(9), 1551-62.

Guo, C., Fischhaber, P. L., Luk-Paszyc, M. J., Masuda, Y., Zhou, J., Kamiya, K., Kisker, C., and Friedberg, E. C. (2003). Mouse Rev1 protein interacts with multiple DNA polymerases involved in translesion DNA synthesis. EMBO Journal, 22(24), 6621-30.

Guo, C., Sonoda, E., Tang, T. S., Parker, J. L., Bielen, A. B., Takeda, S., Ulrich, H. D., and Friedberg, E. C. (2006a). REV1 protein interacts with PCNA: significance of the REV1 BRCT domain *in vitro* and *in vivo*. Molecular Cell, 23(2), 265-71.

Guo, C., Tang, T. S., Bienko, M., Parker, J. L., Bielen, A. B., Sonoda, E., Takeda, S., Ulrich, H. D., Dikic, I., and Friedberg, E. C. (2006b). Ubiquitin-binding motifs in REV1 protein are required for its role in the tolerance of DNA damage. Molecular and Cellular Biology, 26, 8892–8900.

Guo, C., Tang, T.S., Bienko, M., Dikic, I. and Friedberg, E.C. (2008). Requirements for the interaction of mouse Pol κ with ubiquitin, and its biological significance. The Journal of Biological Chemistry, 283, 4658–4664.

Hance, N., Ekstrand, M. I., and Trifunovic, A. (2005). Mitochondrial DNA polymerase gamma is essential for mammalian embryogenesis. Human Molecular Genetics, 14(13), 1775-1783

Haracska, L., Yu, S. L., Johnson, R. E., Prakash, L., and Prakash, S. (2000). Efficient and accurate replication in the presence of 7,8- dihydro-8-oxoguanine by DNA polymerase η. Nature Genetics, 25 458–461.

Haracska, L., Johnson, R. E., Unk, I., Phillips, B., Hurwitz, J., Prakash, L., and Prakash, S. (2001a). Physical and functional interactions of human DNA polymerase η with PCNA. Molecular and Cellular Biology, 21(21), 7199-206.

Haracska, L., Johnson, R. E., Unk, I., Phillips, B., Hurwitz, J., Prakash, L., and Prakash, S. (2001b). Targeting of human DNA polymerase ι to the replication machinery via interaction with PCNA. Proceedings of the National Academy of Sciences, 98(25), 14256-61.

Haracska, L., Unk, I., Johnson, R. E., Phillips, B., Hurwitz, J., Prakash, L., and Prakash, S. (2002a). Stimulation of DNA synthesis activity of human DNA polymerase κ by PCNA. Molecular and Cellular Biology, 22(3), 784-91.

Haracska, L., Prakash, L., and Prakash, S. (2002b). Role of human DNA polymerase κ as an extender in translesion synthesis. Proceedings of the National Academy of Sciences, 99(25), 16000-5.

Hartley, C. L., and McCulloch, R. (2008). *Trypanosoma brucei* BRCA2 acts in antigenic variation and has undergone a recent expansion in BRC repeat number that is important during homologous recombination. Molecular Microbiology, 68(5), 1237-51.

Heller, R. C., and Marians, K. J. (2006). Replication fork reactivation downstream of a blocked nascent leading strand. Nature, 439(7076), 557-62.

Hemsley, A., Arnheim, N., Toney, M. D., Cortopassi, G., and Galas, D. J. (1989). A simple method for site-directed mutagenesis using the polymerase chain reaction. Nucleic Acids Research, 17(16), 6545-51.

Hendel, A., Ziv, O., Gueranger, Q., Geacintov, N., and Livneh, Z. (2008). Reduced efficiency and increased mutagenicity of translesion DNA synthesis across a TT cyclobutane pyrimidine dimer, but not a TT 6-4 photoproduct, in human cells lacking DNA polymerase η. DNA Repair, 7(10), 1636-1646.

Hendel, A., Krijger, P. H., Diamant, N., Goren, Z., Langerak, P., Kim, J., Reissner, T., Lee, K. Y., Geacintov, N. E., Carell, T., Myung, K., Tateishi, S., D'Andrea, A., Jacobs, H., and Livneh, Z. (2011). PCNA ubiquitination is important, but not essential for translesion DNA synthesis in mammalian cells. PLOS Genetics, 7(9), e1002262.

Hines, J. C., and Ray, D. S. (2010). A mitochondrial DNA primase is essential for cell growth and kinetoplast DNA replication in *Trypanosoma brucei*. Molecular and Cellular Biology, 30(6), 1319–1328.

Hines, J. C., and Ray, D. S. (2011). A Second Mitochondrial DNA Primase Is Essential for Cell Growth and Kinetoplast Minicircle DNA Replication in *Trypanosoma brucei*. Eukaryotic Cell, 10(3), 445–454.

Hirota K, Sonoda E, Kawamoto T, Motegi A, Masutani C, Hanaoka F, Szüts D, Iwai S, Sale JE, Lehmann A, Takeda S. (2010) Simultaneous Disruption of Two DNA Polymerases, Polη and Polζ, in Avian DT40 Cells Unmasks the Role of Polη in Cellular Response to Various DNA Lesions. PLoS Genetics, 6(10), e1001151.

Hirumi, H. and Hirumi, K. (1994). Axenic culture of African trypanosome bloodstream forms. Parasitology Today, 10(2), 80-84.

Hoege, C., Pfander, B., Moldovan, G. L., Pyrowolakis, G. and Jentsch, S. (2002). RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. Nature 419, 135-141.

Hoeijmakers, J. H. (2001). Genome maintenance mechanisms for preventing cancer. Nature, 411(6835), 366–374.

Hoffmann, G. R. (1980) Genetic effects of dimethyl sulfate, diethyl sulfate, and related compounds. Mutation Research, 75, 63–129.

Holt, I. J. (2009). Mitochondrial DNA replication and repair: all a flap. Trends in Biochemical Sciences, 34(7), 358–365.

Horn, D., and McCulloch, R. (2010). Molecular mechanisms underlying the control of antigenic variation in African trypanosomes. Current Opinion in Microbiology, 13(6), 700-5.

Hu, J., Guo, L., Wu, K., Liu, B., Lang, S., and Huang, L. (2011). Template-dependent polymerization across discontinuous templates by the heterodimeric primase from the hyperthermophilic archaeon *Sulfolobus solfataricus*. Nucleic Acids Research, 40(8), 3470-3483.

Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2008). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature Protocols, 4(1), 44–57.

Huang, T. T., Nijman, S. M., Mirchandani, K. D., Galardy, P. J., Cohn, M. A., Haas, W., Gygi, S. P., Ploegh, H. L., Bernards, R., and D'Andrea, A. D. (2006). Regulation of monoubiquitinated PCNA by DUB autocleavage. Nature Cell Biology, 8(4), 339-47.

Huberman, J. A., and Riggs, A. D. (1966). Autoradiography of chromosomal DNA fibers from Chinese hamster cells. Proceedings of the National Academy of Sciences, 55(3), 599-606.

Huberman, J. A., and Riggs, A. D. (1968). On the mechanism of DNA replication in mammalian chromosomes. Journal of Molecular Biology, 32(2), 327-41.

Inouye, M., Arnheim, N., and Sternglanz, R. (1973). Bacteriophage T7 lysozyme is an N-acetylmuramyl-L-alanine amidase. The Journal of Biological Chemistry, 248(20), 7247-7252.

Ito, J., and Braithwaite, D. K. (1991). Compilation and alignment of DNA polymerase sequences. Nucleic Acids Research, 19(15), 4045-57.

Ito, J., and Braithwaite, D. K. (1993). Compilation, alignment, and phylogenetic relationships of DNA polymerases. Nucleic Acids Research, 21(4), 787-802.

Ito, N., Nureki, O., Shirouzu, M., Yokoyama, S., and Hanaoka, F. (2003). Crystal structure of the *Pyrococcus horikoshii* DNA primase-UTP complex: implications for the mechanism of primer synthesis. Genes Cells, 8(12), 913-23.

Ivens, A. C., Peacock, C. S., Worthey, E. A., Murphy, L., Aggarwal, G., Berriman, M., Sisk, E., Rajandream, M. A., Adlem, E., Aert, R., Anupama, A., Apostolou, Z., Attipoe, P., Bason, N., Bauser, C., Beck, A., Beverley, S. M., Bianchettin, G., Borzym, K., Bothe, G., Bruschi, C. V., Collins, M., Cadag, E., Ciarloni, L., Clayton, C., Coulson, R. M., Cronin, A., Cruz, A. K., Davies, R. M., De Gaudenzi, J., Dobson, D. E., Duesterhoeft, A., Fazelina, G., Fosker, N., Frasch, A. C., Fraser, A., Fuchs, M., Gabel, C., Goble, A., Goffeau, A., Harris, D., Hertz-Fowler, C., Hilbert, H., Horn, D., Huang, Y., Klages, S., Knights, A., Kube, M., Larke, N., Litvin, L., Lord, A., Louie, T., Marra, M., Masuy, D., Matthews, K., Michaeli, S., Mottram, J. C., Müller-Auer, S., Munden, H., Nelson, S., Norbertczak, H., Oliver, K., O'neil, S., Pentony, M., Pohl, T. M., Price, C., Purnelle, B., Quail, M. A., Rabbinowitsch, E., Reinhardt, R., Rieger, M., Rinta, J., Robben, J., Robertson, L., Ruiz, J. C., Rutter, S., Saunders, D., Schäfer, M., Schein, J., Schwartz, D. C., Seeger, K., Seyler, A., Sharp, S., Shin, H., Sivam, D., Squares, R., Squares, S., Tosato, V., Vogt, C., Volckaert, G., Wambutt, R., Warren, T., Wedler, H., Woodward, J., Zhou, S., Zimmermann, W., Smith, D. F., Blackwell, J. M., Stuart, K. D., Barrell, B., and Myler, P. J. (2005). The genome of the kinetoplastid parasite, *Leishmania major*. Science, 309(5733), 436-42.

Iyer L. M., Koonin, E. V., Leipe D. D., and Aravind L. (2005). Origin and evolution of the archaeo-eukaryotic primase superfamily and related palm-domain proteins: structural insights and new members. Nucleic Acids Research, 33(12), 3875-96.

Jansen, J. G., Langerak, P., Tsaalbi-Shtylik, A., van den Berk, P., Jacobs, H., and de Wind, N. (2006). Strand-biased defect in C/G transversions in hypermutating immunoglobulin genes in Rev1-deficient mice. Journal of Experimental Medicine, 203(2), 319-23.

Jansen, J. G., Tsaalbi-Shtylik, A., Hendriks, G., Gali, H., Hendel, A., Johansson, F., Erixon, K., Livneh, Z., Mullenders, L. H., Haracska, L., and de Wind, N. (2009a). Separate Domains of Rev1 Mediate Two Modes of DNA Damage Bypass in Mammalian Cells. Molecular and Cellular Biology, 29(11), 3113–3123.

Jansen, J. G., Tsaalbi-Shtylik, A., Hendriks, G., Verspuy, J., Gali, H., Haracska, L., and de Wind, N. (2009b). Mammalian polymerase ζ is essential for post-replication repair of UV-induced DNA lesions. DNA Repair, 8(12), 1444–1451.

Jentsch, S., McGrath, J. P. and Varshavsky, A. (1987). The yeast DNA repair gene RAD6 encodes a ubiquitin-conjugating enzyme. Nature, 329, 131-134.

Jiricny, J. (2006). The multifaceted mismatch-repair system. Nature Reviews Molecular Cell Biology, 7(5), 335-46.

Johnson, A., and O'Donnell, M. (2005). Cellular DNA Replicases: Components and Dynamics at the Replication Fork. Annual Review of Biochemistry, 74(1), 283–315.

Johnson, R. E., Prakash, S., and Prakash, L. (1999a) Efficient bypass of a thymine-thymine dimer by yeast DNA polymerase Pol η. Science, 283, 1001–1004.

Johnson, R. E., Kondratick, C. M., Prakash, S., and Prakash, L. (1999b). hRAD30 mutations in the variant form of xeroderma pigmentosum. Science, 285, 263–265.

Johnson, R. E., Washington, M. T., Haracska, L., Prakash, S., and Prakash, L. (2000a). Eukaryotic polymerases ι and ζ act sequentially to bypass DNA lesions. Nature, 406, 1015–1019.

Johnson, R. E., Prakash, S., and Prakash, L. (2000b). The human DINB1 gene encodes the DNA polymerase Pol θ. Proceedings of the National Academy of Sciences, 97(8), 3838-43.

Johnson, R. E., Haracska, L., Prakash, S., and Prakash, L. (2001). Role of DNA polymerase η in the bypass of a (6-4) TT photoproduct. Molecular and Cellular Biology, 21(10), 3558-63.

Johnson, R. E., Yu, S. L., Prakash, S., and Prakash, L. (2003). Yeast DNA polymerase zeta (ζ) is essential for error-free replication past thymine glycol. Genes and Development, 17(1), 77-87.

Johnson, R. E., Prakash, L., and Prakash, S. (2012). Pol31 and Pol32 subunits of yeast DNA polymerase δ are also essential subunits of DNA polymerase ζ. Proceedings of the National Academy of Sciences, 109(31), 12455-60.

Johnson, S. J., and Beese, L. S. (2004). Structures of mismatch replication errors observed in a DNA polymerase. Cell, 116, 803-816.

Joyce, C. M., and Steitz, T. A. (1994) Function and structure relationships in DNA polymerases. Annual Review of Biochemistry, 63, 777–822

Jozwiakowski, S. K., and Connolly, B. A. (2010). A Modified Family-B Archaeal DNA Polymerase with Reverse Transcriptase Activity. ChemBioChem, 12(1), 35–37.

Kaguni, L. S., and Lehman, I. R. (1988). Eukaryotic DNA polymerase-primase: structure, mechanism and function. Biochimica et Biophysica Acta, 950(2), 87-101.

Kaguni, L. S. (2004). DNA polymerase gamma, the mitochondrial replicase. Annual Review of Biochemistry, 73(1), 293–320.

Kannouche, P., Broughton, B. C., Volker, M., Hanaoka, F., Mullenders, L. H. F., and Lehmann, A. R. (2001). Domain structure, localization, and function of DNA polymerase η, defective in xeroderma pigmentosum variant cells. Genes and Development, 15, 158–172.

Kannouche, P., Fernández de Henestrosa, A. R., Coull, B., Vidal, A. E., Gray, C., Zicha, D., Woodgate, R., and Lehmann, A. R. (2003). Localization of DNA polymerases η and ι to the replication machinery is tightly co-ordinated in human cells. The EMBO journal, 22, 1223–1233,

Kannouche, P. L., Wing, J., and Lehmann, A. R. (2004). Interaction of human DNA polymerase η with monoubiquitinated PCNA: a possible mechanism for the polymerase switch in response to DNA damage. Molecular Cell, 14(4), 491-500.

Karras, G. I. and Jentsch, S. (2010). The RAD6 DNA damage tolerance pathway operates uncoupled from the replication fork and is functional beyond S phase. Cell, 141, 255–267.

Kasiviswanathan, R., Gustafson, M. A., Copeland, W. C., and Meyer, J. N. (2011). Human mitochondrial DNA polymerase γ exhibits potential for bypass and mutagenesis at UV-induced cyclobutane thymine dimers. The Journal of Biological Chemistry, 287(12), 9222-9229.

Kato, M., Ito, T., Wagner, G., Richardson, C. C., and Ellenberger, T. (2003). Modular architecture of the bacteriophage T7 primase couples RNA primer synthesis to DNA synthesis. Molecular Cell, 11(5), 1349-60.

Katoh, K., Asimenos, G., and Toh, H. (2009). Multiple alignment of DNA sequences with MAFFT. Methods in Molecular Biology, 537, 39-64.

Kaufmann, W. K. (2010). The human intra-S checkpoint response to UVC-induced DNA damage. Carcinogenesis, 31(5), 751–765.

Kazak, L., Reyes, A., and Holt, I. J. (2012). Minimizing the damage: repair pathways keep mitochondrial DNA intact. Nature Reviews Molecular Cell Biology, 13(10), 659–671.

Keck, J. L., Roche, D. D., Lynch, A. S. and Berger, J. M. (2000). Structure of the RNA polymerase domain of E.coli primase. Science, 287, 2482–2486.

Keck, J. L., and Berger, J. M. (2001). Primus inter pares (first among equals). Nature Structural Biology, 8(1), 2–4.

Kelman, Z., Hurwitz, J., and O'Donnell, M. (1998). Processivity of DNA polymerases: two mechanisms, one goal. Structure, 6(2), 121-5.

Kelmanm, L. M., and Kelman, Z. (2003). Archaea: an archetype for replication initiation studies? Molecular Microbiology, 48(3), 605-15.

Kelner, A. (1949). Effect of visible light on the recovery of *Streptomyces griseus conidia* from ultraviolet irradiation injury. Proceedings of the National Academy of Sciences, 35(2), 73–79.

Kim, N., Mudrak, S. V. and Jinks-Robertson, S. (2011). The dCMP transferase activity of yeast Rev1 is biologically relevant during the bypass of endogenously generated AP sites. DNA Repair, 10(12), 1262-71.

Kirk, B. W., and Kuchta, R. D. (1999). Arg304 of human DNA primase is a key contributor to catalysis and NTP binding: primase and the family X polymerases share significant sequence homology. Biochemistry, 38(24), 7727-36.

Kirouac, K. N. and Ling, H. (2011). Unique active site promotes error-free replication opposite an 8-oxo-guanine lesion by human DNA polymerase ι. Proceedings of the National Academy of Sciences, 108, 3210–3215.

Klenow, H., and Overgaard-Hansen, K. (1970). Proteolytic cleavage of DNA polymerase from *Escherichia coli* B into an exonuclease unit and a polymerase unit. FEBS Letters, 6(1), 25–27.

Kool, E. T. (2002). Active site tightness and substrate fit in DNA replication. Annual Review of Biochemistry, 71:191–219.

Koonin, E. V., Wolf, Y. I., Kondrashov, A. S., and Aravind, L. (2000). Bacterial homologs of the small subunit of eukaryotic DNA primase. Journal of Molecular Microbiology and Biotechnology, 2(4), 509-12.

Kornberg, A., Lehman, I. R., Bessman, M. J., and Simms, E. S. (1956). Enzymic Synthesis of Deoxyribonucleic Acid. Biochimica et Biophysica Acta, 21(1), 197-8.

Kornberg, A. The biological synthesis of deoxyribonucleic acid. Nobel Lecture, December 11, 1959.

Korndörfer, I. P., and Skerra, A. (2002). Improved affinity of engineered streptavidin for the Strep-tag II peptide is due to a fixed open conformation of the lid-like loop at the binding site. Protein Science, 11, 883–893.

Krishna, T. S., Kong, X. P., Gary, S., Burgers, P. M., and Kuriyan, J. (1994). Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. Cell, 79(7), 1233-43.

Kunkel, T., A., and Burgers, P., M. (2008). Dividing the workload at a eukaryotic replication fork. Trends in Cell Biology, 18(11), 521-7.

Kuchta, R. D., and Stengel, G. (2010). Mechanism and evolution of DNA primases. BBA - Proteins and Proteomics, 1804(5), 1180–1189.

Lamb, J. R., Petit-Frère, C., Broughton, B. C., Lehmann, A. R., and Green, M. H. (1989). Inhibition of DNA replication by ionizing radiation is mediated by a trans-acting factor. International Journal of Radiation Biology, 56(2), 125-130.

Lander, E. S. (2011). Initial impact of the sequencing of the human genome. Nature, 470(7333), 187-97.

Lang, G. I., and Murray, A. W. (2011). Mutation rates across budding yeast chromosome VI are correlated with replication timing. Genome Biology and Evolution, 3, 799-811.

Lange, S. S., Takata, K., and Wood, R. D. (2011). DNA polymerases and cancer. Nature Reviews Cancer, 11(2), 96-110.

Lange, S. S., Wittschieben, J. P., and Wood, R. D. (2012). DNA polymerase ζ is required for proliferation of normal mammalian cells. Nucleic Acids Research, 40(10), 4473–4482.

Lao-Sirieix, S. H., and Bell, S. D. (2004). The heterodimeric primase of the hyperthermophilic archaeon *Sulfolobus solfataricus* possesses DNA and RNA primase, polymerase and 3'-terminal nucleotidyl transferase activities. Journal of Molecular Biology, 344(5), 1251-1263.

Lao-Sirieix, S. H., Pellegrini, L., and Bell, S. D. (2005a). The promiscuous primase. Trends in Genetics, 21(10), 568–572.

Lao-Sirieix, S. H., Nookala, R. K., Roversi, P., Bell, S. D., and Pellegrini, L. (2005b). Structure of the heterodimeric core primase. Nature Structural and Molecular Biology, 12(12), 1137-44.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. Bioinformatics, 23(21), 2947–2948.

Larsson, N. G. (2010). Somatic Mitochondrial DNA Mutations in Mammalian Aging. Annual Review of Biochemistry, 79, 683-706.

Lazzaro, F., Novarina, D., Amara, F., Watt, D. L., Stone, J. E., Costanzo, V., Burgers, P. M., Kunkel, T. A., Plevani, P., and Muzi-Falconi, M. (2012). RNase H and Postreplication Repair Protect Cells from Ribonucleotides Incorporated in DNA. Molecular Cell, 45(1), 99–110.

Lee, J. H., Hwang, G. S., and Choi, B. S. (1999). Solution structure of a DNA decamer duplex containing the stable 3' T.G base pair of the pyrimidine(6-4)pyrimidone photoproduct [(6-4) adduct]: implications for the highly specific 3' T --> C transition of the (6-4) adduct. Proceedings of the National Academy of Sciences, 96(12), 6632-6.

Lee, J. H., Bae, S. H., and Choi, B. S. (2000). The Dewar photoproduct of thymidylyl(3'-->5')-thymidine (Dewar product) exhibits mutagenic behavior in accordance with the "A rule". Proceedings of the National Academy of Sciences, 97(9), 4591-6.

Lee, M. G. S., and Van der Ploeg, L. H. T. (1990). Homologous recombination and stable transfection in the parasitic protozoan *Trypanosoma brucei*. Science, 250, 1583–1587.

Lee ,Y. S., Kennedy, W. D., and Yin, Y. W. (2009). Structural insight into processive human mitochondrial DNA synthesis and disease-related polymerase mutations. Cell, 139(2), 312-324.

Lehman, I. R., Bessman, M. J., Simms, E. S., and Kornberg, A. (1958). Enzymatic synthesis of deoxyribonucleic acid. I. Preparation of substrates and partial purification of an enzyme from *Escherichia coli*. The Journal of Biological Chemistry, 233(1), 163-70.

Lehmann, A. R. (1972). Postreplication repair of DNA in ultraviolet-irradiated mammalian cells. Journal of Molecular Biology, 66(3), 319-37.

Lehmann, A. R., Kirk-Bell, S., Arlett, C. F., Paterson, M. C., Lohman, P. H., de Weerd-Kastelein, E. A., and Bootsma, D. (1975). Xeroderma pigmentosum cells with normal levels of excision repair have a defect in DNA synthesis after UV-irradiation. Proceedings of the National Academy of Sciences, 72(1), 219-23.

Lehmann, A. R. (2002). Replication of damaged DNA in mammalian cells: new solutions to an old problem. Mutation Research, 509(1-2), 23–34.

Lehmann, A. R., and Fuchs, R. P. (2006). Gaps and forks in DNA replication: Rediscovering old models. DNA Repair, 5(12), 1495–1498.

Lehmann, A. R. (2011). DNA polymerases and repair synthesis in NER in human cells. DNA Repair, 10(7), 730-3.

Lehmann, A. R., McGibbon, D., and Stefanini, M. (2011). Xeroderma pigmentosum. Orphanet Journal of Rare Diseases, 6, 70.

Lehamnn, A.R. (2011). DNA polymerases and repair synthesis in NER in human cells. DNA repair. 10(7), 730-3.

Leipe, D. D., Aravind, L., and Koonin, E. V. (1999). Did DNA replication evolve twice independently? Nucleic Acids Research, 27(17), 3389-401.

Lemontt, J. F. (1971). Mutants of yeast defective in mutation induced by ultraviolet light. Genetics, 68(1), 21-33.

Letunic, I., and Bork, P. (2011). Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. Nucleic Acids Research, 39, W475-478.

Li, X., and Heyer, W. D. (2008). Homologous recombination in DNA repair and DNA damage tolerance. Cell Research, 18(1), 99–113.

Lim, S. E., Longley, M. J., and Copeland, W. C. (1999). The mitochondrial p55 accessory subunit of human DNA polymerase γ enhances DNA binding, promotes processive DNA synthesis, and confers N-ethylmaleimide resistance. The Journal of Biological Chemistry, 274(53), 38197-203.

Lin, J. R., Zeman, M. K., Chen, J. Y., Yee, M. C., and Cimprich, K. A. (2011). SHPRH and HLTF act in a damage-specific manner to coordinate different forms of postreplication repair and prevent mutagenesis. Molecular Cell, 42(2), 237-49.

Lindahl, T. (1993). Instability and decay of the primary structure of DNA. Nature, 362(6422), 709–715.

Ling H, Boudsocq F, Woodgate R, Yang W. (2001). Crystal structure of a Y-family DNA polymerase in action: a mechanism for error-prone and lesion-bypass replication. Cell, 107(1), 91-102.

Lipps, G., Röther, S., Hart, C., and Krauss, G. (2003). A novel type of replicative enzyme harbouring ATPase, primase and DNA polymerase activity. The EMBO Journal. 22, 2516–2525.

Lipps, G., Weinzierl, A. O., von Scheven, G., Buchen, C., and Cramer, P. (2004) Structure of a bifunctional DNA primase-polymerase. Nature Structural and Molecular Biology, 11(2), 157-62.

Lipscomb, L. A., Peek, M. E., Morningstar, M. L., Verghis, S. M., Miller, E. M., Rich, A., Essigmann, J. M., and Williams, L. D. (1995). X-ray structure of a DNA decamer containing 7,8-dihydro-8-oxoguanine. Proceedings of the National Academy of Sciences, 92(3), 719-23.

Liu, B., Liu, Y., Motyka, S. A., Agbo, E. E., and Englund, P. T. (2005). Fellowship of the rings: the replication of kinetoplast DNA. Trends in Parasitology, 21(8), 363-9.

Liu, L., Komori, K., Ishino, S., Bocquier, A. A., Cann, I. K., Kohda, D., and Ishino, Y. (2001). The archaeal DNA primase – biochemical characterization of the p41-p46 complex from *Pyrococcus furiosus*. Journal of Biological Chemistry, 276, 45484–45490.

Liu, L., Rodriguez-Belmonte, E. M., Mazloum, N., Xie, B., and Lee, M. Y. (2003). Identification of a novel protein, PDIP38, that interacts with the p50 subunit of DNA polymerase δ and proliferating cell nuclear antigen. The Journal of Biological Chemistry, 278, 10041-10047.

Liu, P., Qian, L., Sung, J. S., de Souza-Pinto, N. C., Zheng, L., Bogenhagen, D. F., Bohr, V. A., Wilson, D. M. 3rd, Shen, B., and Demple, B. Removal of oxidative DNA damage via FEN1-dependent long-patch base excision repair in human cell mitochondria. Molecular and Cellular Biology, 28(16), 4975-4987

Loeb, L. A., Springgate, C. F., and Battula, N. (1974). Errors in DNA replication as a basis of malignant changes. Cancer Research, 34(9), 2311-21.

Loeb, L. A. (1991). Mutator phenotype may be required for multistage carcinogenesis. Cancer Research, 51(12), 3075-9.

Lone, S., Townson, S. A., Uljon, S. N., Johnson, R. E., Brahma, A., Nair, D. T., Prakash, S., Prakash, L., and Aggarwal, A. K. (2007). Human DNA polymerase κ encircles DNA: implications for mismatch extension and lesion bypass. Molecular Cell, 25(4), 601-14.

Longhese, M. P., Jovine, L., Plevani, P., and Lucchini, G. (1993). Conditional mutations in the yeast DNA primase genes affect different aspects of DNA metabolism and interactions in the DNA polymerase alpha-primase complex. Genetics, 133(2), 183-191.

Longley, M. J., Prasad, R., Srivastava, D. K., Wilson, S. H., and Copeland, W. C. (1998). Identification of 5'-deoxyribose phosphate lyase activity in human DNA polymerase gamma and its role in mitochondrial base excision repair *in vitro*. Proceedings of the National Academy of Sciences, 95(21), 12244-12248.

Lopes, M., Foiani, M. and Sogo, J. M. (2006). Multiple mechanisms control chromosome integrity after replication fork uncoupling and restart at irreparable UV lesions. Molecular Cell, 21, 15–27.

Lopes Dde, O., Schamber-Reis, B. L., Regis-da-Silva, C. G., Rajão, M. A., Darocha, W. D., Macedo, A. M., Franco, G. R., Nardelli, S. C., Schenkman, S., Hoffmann, J. S., Cazaux, C., Pena, S. D., Teixeira, S. M., and Machado, C. R. (2008). Biochemical studies with DNA polymerase β and DNA polymerase beta-PAK of *Trypanosoma cruzi* suggest the involvement of these proteins in mitochondrial DNA maintenance. DNA Repair, 7(11), 1882-92.

Lowry, O. H., Rosebrough, N. J., Farr, A. L., and Randall, R. J. (1951). Protein measurement with the Folin phenol reagent. The Journal of Biological Chemistry, 193(1), 265-275.

Lucchini, G., Francesconi, S., Foiani, M., Badaracco, G., and Plevani, P. (1987). Yeast DNA polymerase-DNA primase complex; cloning of PRI 1, a single essential gene related to DNA primase activity. The EMBO Journal, 6(3), 737-742.

Lukes, J., Hashimi, H., and Zíková, A. (2005). Unexplained complexity of the mitochondrial genome and transcriptome in kinetoplastid flagellates. Current Genetics, 48(5), 277-99.

MacDougall, C. A., Byun, T. S., Van, C., Yee, M. C., and Cimprich, K. A. (2007). The structural determinants of checkpoint activation. Genes & Development, 21(8), 898–903.

Machado, C. R., Augusto-Pinto, L., McCulloch, R., and Teixeira, S. M. R. (2006). DNA metabolism and genetic diversity in Trypanosomes. Mutation Research, 612(1), 40–57.

Maga, G., Villani, G., Crespan, E., Wimmer, U., Ferrari, E., Bertocci, B., and Hübscher, U. (2007). 8-oxo-guanine bypass by human DNA polymerases in the presence of auxiliary proteins. Nature, 447(7144), 606-608.

Marini, F., Pellicioli, A., Paciotti, V., Lucchini, G., Plevani, P., Stern, D. F., and Foiani, M. (1997). A role for DNA primase in coupling DNA replication to DNA damage response. The EMBO Journal, 16(3), 639–650.

Martínez-Calvillo, S., Vizuet-de-Rueda, J. C., Florencio-Martínez, L. E., Manning-Cela, R. G., and Figueroa-Angulo, E. E. (2010). Gene Expression in Trypanosomatid Parasites. Journal of Biomedicine and Biotechnology, 2010, 1–16.

Masuda, K., Ouchida, R., Li, Y., Gao, X., Mori, H., and Wang, J. Y. (2009). A critical role for REV1 in regulating the induction of C:G transitions and A:T mutations during Ig gene hypermutation. Journal of Immunology, 183(3), 1846-50.

Masutani, C., Araki, M., Yamada, A., Kusumoto, R., Nogimori, T., Maekawa, T., Iwai, S., and Hanaoka, F. (1999a). Xeroderma pigmentosum variant (XP-V) correcting protein from HeLa cells has a thymine dimer bypass DNA polymerase activity. The EMBO Journal, 18(12), 3491-501.

Masutani, C., Kusumoto, R., Yamada, A., Dohmae, N., Yokoi, M., Yuasa, M., Araki, M., Iwai, S., Takio, K., and Hanaoka, F. (1999b). The XPV (xeroderma pigmentosum variant) gene encodes human DNA polymerase η. Nature 399, 700–704.

Masutani, C., Kusumoto, R., Iwai, S., and Hanaoka, F. (2000). Mechanisms of accurate translesion synthesis by human DNA polymerase η. The EMBO Journal, 19(12), 3100-9.

Matsuda, T., Bebenek, K., Masutani, C., Hanaoka, F., and Kunkel, T. A. (2000). Low fidelity DNA synthesis by human DNA polymerase η. Nature, 404(6781), 1011-3.

Matsumoto, T., Eki, T., and Hurwitz, J. (1990). Studies on the initiation and elongation reactions in the simian virus 40 DNA replication system. Proceedings of the National Academy of Sciences, 87(24), 9712-6.

Matsunaga, F., Norais, C., Forterre, P., and Myllykallio, H. (2003). Identification of short 'eukaryotic' Okazaki fragments synthesized from a prokaryotic replication origin. EMBO Reports, 4, 154–158.

Matthews, K. R., Ellis, J. R., and Paterou, A. (2004). Molecular regulation of the life cycle of African trypanosomes. Trends in Parasitology, 20(1), 40-7.

McAteer, K., Jing, Y., Kao, J., Taylor, J. S., and Kennedy, M. A. (1998). Solution-state structure of a DNA dodecamer duplex containing a cis-syn thymine cyclobutane dimer, the major UV photoproduct of DNA. Journal of Molecular Biology, 282(5), 1013-32.

McAuley-Hecht, K. E., Leonard, G. A., Gibson, N. J., Thomson, J. B., Watson, W. P., Hunter, W. N., and Brown, T. (1994). Crystal structure of a DNA duplex containing 8-hydroxydeoxyguanine-adenine base pairs. Biochemistry, 33(34), 10266-10270.

McCulloch, R., and Barry, J. D. (1999). A role for RAD51 and homologous recombination in Trypanosoma brucei antigenic variation. Genes and Development, 13(21), 2875-88.

McCulloch, S. D., Kokoska, R. J., Masutani, C., Iwai, S., Hanaoka, F., and Kunkel, T. A. (2004). Preferential cis-syn thymine dimer bypass by DNA polymerase η occurs with biased fidelity. Nature, 428(6978), 97-100.

McCulloch, S. D., and Kunkel, T. A. (2008). The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. Cell Research, 18(1), 148–161.

McCulloch, S. D., Kokoska, R. J., Garg, P., Burgers, P. M., and Kunkel, T. A. (2009). The efficiency and fidelity of 8-oxo-guanine bypass by DNA polymerases δ and η. Nucleic Acids Research, 37(9), 2830-40.

McDonald, J. P., Rapić-Otrin, V., Epstein, J. A., Broughton, B. C., Wang, X., Lehmann, A. R., Wolgemuth, D. J., and Woodgate, R. (1999). Novel human and mouse homologs of Saccharomyces cerevisiae DNA polymerase η. Genomics, 60(1), 20-30.

McDonald, J. P., Mead, S., and Woodgate, R. (2006). Controlling the subcellular localization of DNA polymerases ι and η via interactions with ubiquitin. The EMBO Journal, 25(12), 2847-55.

MacDougall, C. A., Byun, T. S., Van, C., Yee, M. C., and Cimprich, K. A. (2007) The structural determinants of checkpoint activation. Genes and Development, 21(8), 898-903.

McIlwraith, M. J., Vaisman, A., Liu, Y., Fanning, E., Woodgate, R., and West, S. C. (2005). Human DNA polymerase η promotes DNA synthesis from strand invasion intermediates of homologous recombination. Molecular Cell, 20 783–792.

McIntyre, J., Vidal, A. E., McLenigan, M. P., Bomar, M. G., Curti, E., McDonald, J. P., Plosky, B. S., Ohashi, E., and Woodgate, R. (2012). Ubiquitin mediates the physical and functional

interaction between human DNA polymerases η and ι. Nucleic Acids Research. Epub ahead of print, doi: 10.1093/nar/gks1277.

McNally, K., Neal, J. A., McManus, T. P., McCormick, J. J., and Maher, V. M. (2008). hRev7, putative subunit of hPol ζ, plays a critical role in survival, induction of mutations, and progression through S-phase, of UV((254nm))-irradiated human fibroblasts. DNA Repair, 7(4), 597-604.

Méchali, M. (2010). Eukaryotic DNA replication origins: many choices for appropriate answers. Nature Reviews Molecular Cell Biology, (10), 728-38.

Melendy, T., and Stillman, B. (1991). Purification of DNA polymerase δ as an essential simian virus 40 DNA replication factor. The Journal of Biological Chemistry, 266(3), 1942-9.

Michael, W. M., Ott, R., Fanning, E., and Newport, J. (2000). Activation of the DNA replication checkpoint through RNA synthesis by primase. Science, 289(5487), 2133-2137.

Mitchell, D. L., Brash, D. E., and Nairn, R. S. (1990). Rapid repair kinetics of pyrimidine (6-4) pyrimidone photoproducts in human cells are due to excision rather than conformational change. Nucleic Acids Research, 18(4), 963-971

Mizuno, T., Okamoto, T., Yokoi, M., Izumi, M., Kobayashi, A., Hachiya, T., Tamai, K., Inoue, T., Hanaoka, F. (1996) Identification of the nuclear localization signal of mouse DNA primase: nuclear transport of p46 subunit is facilitated by interaction with p54 subunit. Journal of Cell Science, 109, 2627–2636.

Moffatt, B. A., and Studier, F. W. (1987). T7 lysozyme inhibits transcription by T7 RNA polymerase. Cell, 49(2), 221-227.

Moore, E. C. (1969). The effects of ferrous ion and dithioerythritol on inhibition by hydroxyruea of ribonucleotide reductase. Cancer Research, 29(2), 291-295.

Morrison, A., Araki, H., Clark, A. B., Hamatake, R. K., and Sugino, A. (1990) A third essential DNA polymerase in *S. cerevisiae*. Cell, 62(6), 1143-51.

Murakami, Y., Eki, T., and Hurwitz, J. (1992). Studies on the initiation of simian virus 40 replication in vitro: RNA primer synthesis and its elongation. Proceedings of the National Academy of Sciences, 89(3), 952-6.

Murakumo, Y., Ogura, Y., Ishii, H., Numata, S., Ichihara, M., Croce, C. M., Fishel, R., and Takahashi, M. (2001). Interactions in the error-prone postreplication repair proteins hREV1, hREV3, and hREV7. The Journal of Biological Chemistry, 276(38), 35644-51.

Muzi-Falconi, M., Giannattasio, M., Foiani, M., and Plevani, P. (2003). The DNA Polymerase α-Primase Complex: Multiple Functions and Interactions. The Scientific World Journal, 3, 21–33.

Nair, D. T., Johnson, R. E., Prakash, L., Prakash, S., and Aggarwal, A. K. (2004). Replication by human DNA polymerase ι occurs by Hoogsteen base-pairing. Nature, 430(6997), 377-80.

Nair, D. T., Johnson, R. E., Prakash, L., Prakash, S., and Aggarwal, A. K. (2005). Rev1 employs a novel mechanism of DNA synthesis using a protein template. Science, 309(5744), 2219-22.

Nasheuer, H. P., Moore, A., Wahl, A. F., and Wang, T. S. (1991). Cell cycle-dependent phosphorylation of human DNA polymerase alpha. The Journal of Biological Chemistry, 266(12), 7893-903.

Nelson, J. R., Lawrence, C. W., and Hinkle, D. C. (1996a). Thymine-thymine dimer bypass by yeast DNA polymerase ζ. Science, 272(5268), 1646-9.

Nelson, J. R., Lawrence, C. W., and Hinkle, D. C. (1996b). Deoxycytidyl transferase activity of yeast REV1 protein. Nature, 382(6593), 729-31.

Nelson, J. R., Gibbs, P. E., Nowicka, A. M., Hinkle, D. C., and Lawrence, C. W. (2000). Evidence for a second function for Saccharomyces cerevisiae Rev1p. Molecular Microbiology, 37(3), 549-54.

Nick McElhinny, S. A., Gordenin, D. A., Stith, C. M., Burgers, P. M., and Kunkel, T. A. (2008). Division of labor at the eukaryotic replication fork. Molecular Cell, 30(2), 137-44.

Nick McElhinny, S. A., Watts, B. E., Kumar, D., Watt, D. L., Lundström, E. B., Burgers, P. M., Johansson, E., Chabes, A., and Kunkel, T. A. (2010). Abundant ribonucleotide incorporation into DNA by yeast replicative polymerases. Proceedings of the National Academy of Sciences, 107(11), 4949-54.

Niimi, A., Brown, S., Sabbioneda, S., Kannouche, P. L., Scott, A., Yasui, A., Green, C. M., and Lehmann, A. R. (2008). Regulation of proliferating cell nuclear antigen ubiquitination in mammalian cells. Proceedings of the National Academy of Sciences, 105(42), 16125-30.

Niimi, A., Chambers, A. L., Downs, J. A., and Lehmann, A. R. (2012). A role for chromatin remodellers in replication of damaged DNA. Nucleic Acids Research, 40(15), 7393-403.

Notredame, C., Higgins, D. G., and Heringa, J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. Journal of Molecular Biology, 302(1), 205-17.

Oakley, G. G., and Patrick, S. M. (2010). Replication protein A: directing traffic at the intersection of replication and repair. Frontiers in Bioscience, 15, 883-900.

Ochman, H., Gerber, A. S., and Hartl, D. L. (1988). Genetic applications of an inverse polymerase chain reaction. Genetics, 120(3), 621-3.

Ogi, T., Kato, T. Jr, Kato, T., and Ohmori, H. (1999). Mutation enhancement by DINB1, a mammalian homologue of the *Escherichia coli* mutagenesis protein dinB. Genes to Cells, 4(11), 607-18.

Ogi, T., Shinkai, Y., Tanaka, K., and Ohmori, H. (2002). Pol κ protects mammalian cells against the lethal and mutagenic effects of benzo[a]pyrene. Proceedings of the National Academy of Sciences, 99(24), 15548-53.

Ogi, T., Kannouche, P., and Lehmann, A. R. (2005). Localisation of human Y-family DNA polymerase κ: relationship to PCNA foci. Journal of Cell Science, 118, 129–136.

Ogi, T., and Lehmann, A. R. (2006). The Y-family DNA polymerase kappa (pol κ) functions in mammalian nucleotide-excision repair. Nature Cell Biology, 8(6), 640-2.

Ogi, T., Limsirichaikul, S., Overmeer, R. M., Volker, M., Takenaka, K., Cloney, R., Nakazawa, Y., Niimi, A., Miki, Y., Jaspers, N. G., Mullenders, L. H., Yamashita, S., Fousteri, M. I., and Lehmann, A. R. (2010). Three DNA polymerases, recruited by different mechanisms, carry out NER repair synthesis in human cells. Molecular Cell, 37(5), 714-27.

Ohashi, E., Ogi, T., Kusumoto, R., Iwai, S., Masutani, C., Hanaoka, F., and Ohmori, H. (2000). Error-prone bypass of certain DNA lesions by the human DNA polymerase κ. Genes and Development, 14(13), 1589-94.

Ohashi, E., Murakumo, Y., Kanjo, N., Akagi, J., Masutani, C., Hanaoka, F., and Ohmori, H. (2004). Interaction of hREV1 with three human Y-family DNA polymerases. Genes Cells, 9(6), 523-31.

Ohkumo, T., Kondo, Y., Yokoi, M., Tsukamoto, T., Yamada, A., Sugimoto, T., Kanao, R., Higashi, Y., Kondoh, H., Tatematsu, M., Masutani, C., and Hanaoka, F. (2006). UV-B radiation induces epithelial tumors in mice lacking DNA polymerase η and mesenchymal tumors in mice deficient for DNA polymerase ι. Molecular and Cellular Biology, 26(20), 7696-706.

Ohmori, H., Friedberg, E. C., Fuchs, R. P., Goodman, M. F., Hanaoka, F., Hinkle, D., Kunkel, T. A., Lawrence, C. W., Livneh, Z., Nohmi, T., Prakash, L., Prakash, S., Todo, T., Walker, G. C., Wang, Z., and Woodgate, R. (2001). The Y-family of DNA polymerases. Molecular Cell, 8(1), 7-8.

Okazaki, R., Okazaki, T., Sakabe, K., Sugimoto, K., and Sugino, A. (1968) Mechanism of DNA chain growth. I. Possible discontinuity and unusual secondary structure of newly synthesized chains. Proceedings of the National Academy of Sciences, (2), 598-605.

Ollis, D. L., Brick, P., Hamlin, R., Xuong, N. G., and Steitz, T. A. (1985). Structure of large fragment of *Escherichia coli* DNA polymerase I complexed with dTMP. Nature, 313 (6005), 762–766.

Olson, E., Nievera, C. J., Klimovich, V., Fanning, E., and Wu, X. (2006). RPA2 is a direct downstream target for ATR to regulate the S-phase checkpoint. Journal of Biological Chemistry, 281(51), 39517-39533.

Otsuka, C., Kunitomi, N., Iwai, S., Loakes, D., and Negishi, K. (2005). Roles of the polymerase and BRCT domains of Rev1 protein in translesion DNA synthesis in yeast *in vivo*. Mutation Research, 578(1-2), 79-87.

Otterlei, M., Warbrick, E., Nagelhus, T. A., Haug, T., Slupphaug, G., Akbari, M., Aas, P. A., Steinsbekk, K., Bakke, O., and Krokan, H. E. (1999). Post-replicative base excision repair in replication foci. The EMBO Journal, 18(13), 3834-3844.

Passos-Silva, D. G., Rajão, M. A., Nascimento De Aguiar, P. H., Vieira-Da-Rocha, J. P., Machado, C. R., and Furtado, C. (2010). Overview of DNA Repair in *Trypanosoma cruzi*, *Trypanosoma brucei*, and *Leishmania major*. Journal of Nucleic Acids, 840768.

Pathania, S., Nguyen, J., Hill, S. J., Scully, R., Adelmant, G. O., Marto, J. A., Feunteun, J., and Livingston, D. M. (2011). BRCA1 is required for postreplication repair after UV-induced DNA damage. Molecular Cell, 44(2), 235-251.

Pavlov, Y. I., and Shcherbakova, P. V. (2010). DNA polymerases at the eukaryotic fork-20 years later. Mutation Research, 685(1-2), 45-53.

Pelletier, H., Sawaya, M. R., Kumar, A., Wilson, S. H., and Kraut, J. (1994). Structures of ternary complexes of rat DNA polymerase β, a DNA template-primer, and ddCTP. Science, 264(5167), 1891-903.

Petta, T. B., Nakajima, S., Zlatanou, A., Despras, E., Couve-Privat, S., Ishchenko, A., Sarasin, A., Yasui, A., and Kannouche, P. (2008). Human DNA polymerase ι protects cells against oxidative stress. The EMBO Journal, 27(21), 2883-95.

Phillips, D. H. (1983). Fifty years of benzo(a)pyrene. Nature, 303(5917), 468-72.

Pitcher, R. S., Tonkin, L. M., Green, A. J., and Doherty, A. J. (2005). Domain Structure of a NHEJ DNA Repair Ligase from Mycobacterium tuberculosis. Journal of Molecular Biology, 351(3), 531–544.

Pitcher, R. S., Tonkin, L. M., Daley, J. M., Palmbos, P. L., Green, A. J., Velting, T. L., Brzostek, A., Korycka-Machala, M., Cresawn, S., Dziadek, J., Hatfull, G. F., Wilson, T. E., and Doherty, A. J. (2006). Mycobacteriophage exploit NHEJ to facilitate genome circularization. Molecular Cell, 23(5), 743-8.

Pitcher, R. S., Brissett, N. C., Picher, A. J., Andrade, P., Juarez, R., Thompson, D., Fox, G. C., Blanco, L., and Doherty, A. J. (2007a). Structure and Function of a Mycobacterial NHEJ DNA Repair Polymerase. Journal of Molecular Biology, 366(2), 391–405.

Pitcher, R. S., Brissett, N. C., and Doherty, A. J. (2007b). Nonhomologous End-Joining in Bacteria: A Microbial Perspective. Annual Review of Microbiology, 61(1), 259–282.

Plosky, B. S., Vidal, A. E., Fernandez de Henestrosa, A. R., McLenigan, M. P., McDonald, J. P., Mead, S. and Woodgate, R. (2006). Controlling the subcellular localization of DNA polymerases iota and eta via interactions with ubiquitin. EMBO Journal, 25, 2847–2855.

Plosky, B. S., Frank, E. G., Berry, D. A., Vennall, G. P., McDonald, J. P., and Woodgate, R. (2008). Eukaryotic Y-family polymerases bypass a 3-methyl-2'-deoxyadenosine analog in vitro and methyl methanesulfonate-induced DNA damage *in vivo*. Nucleic Acids Research, 36(7), 2152–2162.

Podobnik, M., McInerney, P., O'Donnell, M., and Kuriyan, J. (2000). A TOPRIM domain in the crystal structure of the catalytic core of Escherichia coli primase confirms a structural link to DNA topoisomerases. Journal of Molecular Biology, 300(2), 353-62.

Prakash, L. (1981). Characterization of postreplication repair in *Saccharomyces cerevisiae* and effects of rad6, rad18, rev3 and rad52 mutations. Molecular and General Genetics, 184(3), 471-8.

Prasad, R., Bebenek, K., Hou, E., Shock, D. D., Beard, W. A., Woodgate, R., Kunkel, T. A., and Wilson, S. H. (2003). Localization of the deoxyribose phosphate lyase active site in human DNA polymerase ι by controlled proteolysis. The Journal of Biological Chemistry, 278(32), 29649-54.

Prelich, G., Kostura, M., Marshak, D. R., Mathews, M. B., and Stillman, B. (1987). The cell-cycle regulated proliferating cell nuclear antigen is required for SV40 DNA replication in vitro. Nature, 326(6112), 471-5.

Proudfoot, C., and McCulloch, R. (2005). Distinct roles for two RAD51-related genes in *Trypanosoma brucei* antigenic variation. Nucleic Acids Research, 33(21), 6906-19.

Pursell, Z. F., Isoz, I., Lundström, E. B., Johansson, E., and Kunkel, T. A. (2007). Yeast DNA polymerase epsilon participates in leading-strand DNA replication. Science, 317(5834),127-30.

Rajão, M. A., Passos-Silva, D. G., DaRocha, W. D., Franco, G. R., Macedo, A. M., Pena, S. D. J., Teixeira, S. M., and Machado, C. R. (2009). DNA polymerase κ from *Trypanosoma cruzi* localizes to the mitochondria, bypasses 8-oxoguanine lesions and performs DNA synthesis in a recombination intermediate. Molecular Microbiology, 71(1), 185–197.

Rasmussen, R. E., and Painter, R. B. (1964). Evidence for repair of ultra-violet damaged deoxyribonucleic acid in cultured mammalian cells. Nature, 203, 1360-2.

Rastogi, R. P., Richa, Kumar, A., Tyagi, M. B., and Sinha, R. P. (2010). Molecular Mechanisms of Ultraviolet Radiation-Induced DNA Damage and Repair. Journal of Nucleic Acids, 1–32.

Redmond, S., Vadivelu, J., and Field, M. C. (2003). RNAit: an automated web-based tool for the selection of RNAi targets in *Trypanosoma brucei*. Molecular and Biochemical Parasitology, 128, 115–118.

Reijns, M. A., Rabe, B., Rigby, R. E., Mill, P., Astell, K. R., Lettice, L. A., Boyle, S., Leitch, A., Keighren, M., Kilanowski, F., Devenney, P. S., Sexton, D., Grimes, G., Holt, I. J., Hill, R. E., Taylor, M. S., Lawson, K. A., Dorin, J. R., and Jackson, A. P. (2012). Enzymatic removal of ribonucleotides from DNA is essential for mammalian genome integrity and development. Cell, 149(5), 1008-22.

Rey, L., Sidorova, J. M., Puget, N., Boudsocq, F., Biard, D. S., Monnat, R. J. Jr, Cazaux, C., and Hoffmann, J. S. (2009). Human DNA polymerase η is required for common fragile site stability during unperturbed DNA replication. Molecular and Cell Biology, 29(12), 3344-54.

Reyes, A., He, J., Mao, C. C., Bailey, L. J., Di Re, M., Sembongi, H., Kazak, L., Dzionek, K., Holmes, J. B., Cluett, T. J., Harbour, M. E., Fearnley, I. M., Crouch, R. J., Conti, M. A., Adelstein, R. S., Walker, J. E., and Holt, I. J. (2011). Actin and myosin contribute to mammalian mitochondrial DNA maintenance. Nucleic Acids Research 39(12), 5098-5108.

Robison, J. G., Elliott, J., Dixon, K., and Oakley, G. G. (2004). Replication Protein A and the Mre11-Rad50-Nbs1 Complex Co-localize and Interact at Sites of Stalled Replication Forks. The Journal of Biological Vhemistry, 279(33), 34802–34810.

Ross, A. L. and Sale, J. E. (2006). The catalytic activity of REV1 is employed during immunoglobulin gene diversification in DT40. Molecular Immunology, 43, 1587–1594.

Rothwell, P. J., and Waksman, G. (2005). Structure and Mechanism of DNA Polymerases. Advances in Protein Chemistry, 71, 401–440.

Rowen, L., and Kornberg, A. (1978). Primase, the dnaG protein of *Escherichia coli*. An enzyme which starts DNA chains. Journal of Biological Chemistry, 253(3), 758-64.

Rupp, W. D., and Howard-Flanders, P. (1968). Discontinuities in the DNA synthesized in an excision-defective strain of Escherichia coli following ultraviolet irradiation. Journal of Molecular Biology, 31(2), 291-304.

Rymer, R. U., Solorio, F. A., Tehranchi, A. K., Chu, C., Corn, J. E., Keck, J. L., Wang, J. D., and Berger, J. M. (2012). Binding mechanism of metal·NTP substrates and stringent-response alarmones to bacterial DnaG-type primases. Structure, 20(9), 1478-89.

Sabbioneda, S., Green, C. M., Bienko, M., Kannouche, P., Dikic, I., and Lehmann, A. R. (2009). Ubiquitin-binding motif of human DNA polymerase η is required for correct localization. Proceedings of the National Academy of Sciences, 106(8), E20.

Sakabe, K., and Okazaki, R. (1966). A unique property of the replicating region of chromosomal DNA. Biochimica et Biophysica Acta, 129(3), 651-4.

Sale, J. E. (2012). Competition, collaboration and coordination - determining how cells bypass DNA damage. Journal of Cell Science, 125(7), 1633–1643.

Sale, J. E., Lehmann, A. R., and Woodgate, R. (2012). Y-family DNA polymerases and their role in tolerance of cellular DNA damage. Nature Reviews Molecular Cell Biology, 13(3), 141–152.

Sambrook, J. and Russell, D. W. (2001). Molecular Cloning: A Laboratory Manual. New York, Cold Spring Harbour Laboratory Press.

Santamaria, D., Viguera, E., Martinez-Robles, M. L., Hyrien, O., Hernandez, P., Krimer, D. B. and Schvartzman, J. B. (2000). Bi-directional replication and random termination. Nucleic Acids Research. 28, 2099–2107.

Sarkies, P., Reams, C., Simpson, L. J., and Sale, J. E. (2010). Epigenetic instability due to defective replication of structured DNA. Molecular Cell, 40(5), 703-13.

Saxowsky, T. T., Choudhary, G., Klingbeil, M. M., and Englund, P. T. (2003). *Trypanosoma brucei* has two distinct mitochondrial DNA polymerase β enzymes. Journal of Biological Chemistry, 278(49), 49095-101.

Sayers, E. W., Barrett, T., Benson, D. A., Bryant, S. H., Canese, K., Chetvernin, V., Church, D. M., DiCuccio, M., Edgar, R., Federhen, S., Feolo, M., Geer, L. Y., Helmberg, W., Kapustin, Y., Landsman, D., Lipman, D. J., Madden, T. L., Maglott, D. R., Miller, V., Mizrachi, I., Ostell, J., Pruitt, K. D., Schuler, G. D., Sequeira, E., Sherry, S. T., Shumway, M., Sirotkin, K., Souvorov, A., Starchenko, G., Tatusova, T. A., Wagner, L., Yaschenko, E., and Ye, J. (2009). Database resources of the National Center for Biotechnology Information. Nucleic Acids Research, 37(Database issue), D5-15.

Scherzinger, E., Lanka, E., Morelli, G., Seiffert, D., and Yuki, A. (1977). Bacteriophage-T7-induced DNA-priming protein. A novel enzyme involved in DNA replication. European Journal of Biochemistry, 72(3), 543-58.

Schamber-Reis, B. L., Nardelli, S., Régis-Silva, C. G., Campos, P. C., Cerqueira, P. G., Lima, S. A., Franco, G. R., Macedo, A. M., Pena, S. D., Cazaux, C., Hoffmann, J. S., Motta, M. C., Schenkman, S., Teixeira, S. M., and Machado, C. R. (2012). DNA polymerase β from *Trypanosoma cruzi* is involved in kinetoplast DNA replication and repair of oxidative lesions. Molecular and Biochemical Parasitology, 183(2), 122-31.

Schmidt, T. G., Koepke, J., Frank, R., and Skerra, A. (1996). Interaction Between the Strep-tag Affinity Peptide and its Cognate Target, Streptavidin. Journal of Molecular Biology, 255, 753–766.

Schmutz, V., Janel-Bintz, R., Wagner, J., Biard, D., Shiomi, N., Fuchs, R. P., and Cordonnier, A. M. (2010). Role of the ubiquitin-binding domain of Polη in Rad18-independent translesion DNA synthesis in human cell extracts. Nucleic Acids Research, 38(19), 6456-65.

Schneider, A., Smith, R. W., Kautz, A. R., Weisshart, K., Grosse, F., and Nasheuer, H. P. (1998). Primase activity of human DNA polymerase alpha-primase. Divalent cations stabilize the enzyme activity of the p48 subunit. The Journal of Biological Chemistry, 273(34), 21608-21615.

Schub, O., Rohaly, G., Smith, R. W., Schneider, A., Dehde, S., Dornreiter, I., Nasheuer, H. P. (2001). Multiple phosphorylation sites of DNA polymerase alpha-primase cooperate to regulate the initiation of DNA replication *in vitro*. The Journal of Biological Chemistry, 276(41), 38076-38083.

Setlow, R. B., Swenson, P. A., and Carrier, W. L. (1963). Thymine dimers and inhibition of DNA synthesis by ultraviolet irradiation of cells. Science, 142(3598), 1464-6.

Setlow, R. B., and Carrier, W. L. (1964). The disappearance of thymine dimers from DNA: an error-correcting mechanism, Proceedings of the National Academy of Sciences, 51, 226–231.

Sharma, S., Hicks, J. K., Chute, C. L., Brennan, J. R., Ahn, J. Y., Glover, T. W., and Canman, C. E. (2012). REV1 and polymerase ζ facilitate homologous recombination repair. Nucleic Acids Research, 40(2), 682-91.

Sheaff, R. J., and Kuchta, R. D. (1993). Mechanism of calf thymus DNA primase: slow initiation, rapid polymerization, and intelligent termination. Biochemistry, 32(12), 3027-37.

Siegel, T. N., Hekstra, D. R., and Cross, G. A. (2008). Analysis of the *Trypanosoma brucei* cell cycle by quantitative DAPI imaging. Molecular and Biochemical Parasitology, 60(2), 171-174.

Silverstein, T. D., Johnson, R. E., Jain, R., Prakash, L., Prakash, S., and Aggarwal, A. K. (2010). Structural basis for the suppression of skin cancers by DNA polymerase η. Nature, 465(7301), 1039-43.

Simmons, J. M., Koslowsky, D. J., and Hausinger, R. P. (2012). Characterization of a Trypanosoma brucei Alkb homolog capable of repairing alkylated DNA. Experimental Parasitology, 131(1), 92-100.

Simpson, A. G., Stevens, J. R., and Lukes, J. (2006). The evolution and diversity of kinetoplastid flagellates. Trends in Parasitology, 22(4), 168-74.

Soria, G., Podhajcer, O., Prives, C. and Gottifredi, V. (2006). P21 (Cip1/WAF1) downregulation is required for efficient PCNA ubiquitination after UV irradiation. Oncogene, 25, 2829–2838.

Stamatoyannopoulos, J. A., Adzhubei, I., Thurman, R. E., Kryukov, G. V., Mirkin, S. M., Sunyaev, S. R. (2009). Human mutation rate associated with DNA replication timing. Nature Genetics, 41(4), 393-395.

Steitz, T. A. (1993). DNA- and RNA-dependent DNA polymerases. Current Opinion in Structural Biology, 3, 31–38.

Steitz, T. A., Smerdon, S. J., Jäger, J., and Joyce, C. M. (1994). A unified polymerase mechanism for nonhomologous DNA and RNA polymerases. Science, 23, 266(5193), 2022-5.

Steitz, T. A. (1999). DNA polymerases: structural diversity and common mechanisms. The Journal of Biological Chemistry, 274(25), 17395–17398.

Stelter, P., and Ulrich, H. D. (2003). Control of spontaneous and damage-induced mutagenesis by SUMO and ubiquitin conjugation. Nature, 425(6954), 188-91.

Studier, F. W. (1991). Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. Journal of Molecular Biology, 219(1), 37-44.

Sugino, A., Hirose, S., and Okazaki, R. (1972). RNA-linked nascent DNA fragments in *Escherichia coli*. Proceedings of the National Academy of Sciences, 69(7), 1863-7.

Sugino, A. (1995). Yeast DNA polymerases and their role at the replication fork. Trends in Biochemical Sciences, 20(8), 319-323.

Sundin, O. and Varshavsky, A. (1980). Terminal stages of SV40 DNA replication proceed via multiply intertwined catenated dimers. Cell, 21, 103–114.

Svoboda, D. L., and Vos, J. M. (1995). Differential replication of a single, UV-induced lesion in the leading or lagging strand by a human cell extract: fork uncoupling or gap formation. Proceedings of the National Academy of Sciences, 92(26), 11975-9.

Szüts, D., Marcus, A. P., Himoto, M., Iwai, S., and Sale, J. E. (2008). REV1 restrains DNA polymerase ζ to ensure frame fidelity during translesion synthesis of UV photoproducts *in vivo*. Nucleic Acids Research, 36(21), 6767-6780.

Taladriz, S., Hanke, T., Ramiro, M. J., García-Díaz, M., García De Lacoba, M., Blanco, L., and Larraga, V. (2001). Nuclear DNA polymerase β from Leishmania infantum. Cloning, molecular analysis and developmental regulation. Nucleic Acids Research, 29(18), 3822-34.

Temviriyanukul, P., van Hees-Stuivenberg, S., Delbos, F., Jacobs, H., de Wind, N., and Jansen, J. G. (2012). Temporally distinct translesion synthesis pathways for ultraviolet light-induced photoproducts in the mammalian genome. DNA Repair, 11(6),550-8.

Ten Asbroek, A. L. M. A., Ouellette, M., and Borst, P. (1990). Targeted insertion of the neomycin phosphotransferase gene into the tubulin gene cluster of *Trypanosoma brucei*. Nature, 348, 174–175.

Thomas, D. C., Roberts, J. D., Sabatino, R. D., Myers, T. W., Tan, C. K., Downey, K. M., So, A. G., Bambara, R. A., and Kunkel, T. A. (1991). Fidelity of mammalian DNA replication and replicative DNA polymerases. Biochemistry, 30(51),11751-9.

Tiengwe, C., Marcello, L., Farr, H., Gadelha, C., Burchmore, R., Barry, J. D., Bell, S. D., and McCulloch, R. (2012a). Identification of ORC1/CDC6-interacting factors in *Trypanosoma brucei* reveals critical features of origin recognition complex architecture. PLoS One, 7(3), e32674.

Tiengwe, C., Marcello, L., Farr, H., Dickens, N., Kelly, S., Swiderski, M., Vaughan, D., Gull, K., Barry, J. D., Bell, S. D., and McCulloch, R. (2012b). Genome-wide analysis reveals extensive functional interaction between DNA replication initiation and transcription in the genome of *Trypanosoma brucei*. Cell Reports, 2(1), 185-97.

Tissier, A., McDonald, J. P., Frank, E. G., and Woodgate, R. (2000a). Pol ι, a remarkably error-prone human DNA polymerase. Genes and Development, 14(13), 1642-50.

Tissier, A., Frank, E. G., McDonald, J. P., Iwai, S., Hanaoka, F., and Woodgate, R. (2000b). Misinsertion and bypass of thymine-thymine dimers by human DNA polymerase iota. The EMBO Journal, 19(19), 5259-66.

Tissier, A., Kannouche, P., Reck, M, P., Lehmann, A. R., Fuchs, R. P. P., and Cordonnier, A. (2004). Co-localization in replication foci and interaction of human Y- family members, DNA polymerase pol η and REVl protein. DNA repair, 3, 1503–1514.

Tissier, A. S., Janel-Bintz, R., Coulon, S., Klaile, E., Kannouche, P., Fuchs, R. P., and Cordonnier, A. M. (2010). Crosstalk between replicative and translesional DNA polymerases: PDIP38 interacts directly with Pol η. DNA repair, 9(8), 922–928.

Torres-Ramos, C. A., Yoder, B. L., Burgers, P. M., Prakash, S. and Prakash, L. (1996). Requirement of proliferating cell nuclear antigen in RAD6-dependent postreplicational DNA repair. Proceedings of the National Academy of Sciences, 93, 9676-9681.

Trenaman, A., Hartley, C., Prorocic, M., Passos-Silva, D. G., Hoek, M. V., Nechyporuk-Zloy, V., Machado, C. R., and McCulloch, R. (2012). *Trypanosoma brucei* BRCA2 acts in a life cycle-specific genome stability process and dictates BRC repeat number-dependent RAD51 subnuclear dynamics. Nucleic Acids Research. Epub ahead of print, doi: 10.1093/nar/gks1192.

Trifunovic, A., Wredenberg, A., Falkenberg, M., Spelbrink, J. N., Rovio, A. T., Bruder, C. E., Bohlooly-Y, M., Gidlöf, S., Oldfors, A., Wibom, R., Törnell, J., Jacobs, H. T., and Larsson, N. G. (2004). Premature ageing in mice expressing defective mitochondrial DNA polymerase. Nature, 429(6990), 417-23.

Trincao, J., Johnson, R. E., Escalante, C. R., Prakash, S., Prakash, L., and Aggarwal, A. K. (2001). Structure of the catalytic core of S. cerevisiae DNA polymerase η: implications for translesion DNA synthesis. Molecular Cell, 8(2), 417-26.

Tseng, B. Y., and Ahlem, C. N. (1984). Mouse primase initiation sites in the origin region of SV40. Proceedings of the National Academy of Sciences, 81(8), 2342-6.

Tsurimoto, T., and Stillman, B. (1989). Purification of a cellular replication factor, RF-C, that is required for coordinated synthesis of leading and lagging strands during simian virus 40 DNA replication in vitro. Molecular and Cellular Biology, 9(2), 609-19.

Tsurimoto, T., Melendy, T., and Stillman, B. (1990) Sequential initiation of lagging and leading strand synthesis by two different polymerase complexes at the SV40 DNA replication origin. Nature, 346(6284), 534-9.

Tsurimoto, T., and Stillman, B. (1991a). Replication factors required for SV40 DNA replication in vitro. I. DNA structure-specific recognition of a primer-template junction by eukaryotic DNA polymerases and their accessory proteins. Journal of Biological Chemistry, 266(3),1950-60.

Tsurimoto, T., and Stillman, B. (1991b). Replication factors required for SV40 DNA replication in vitro. II. Switching of DNA polymerase α and δ during initiation of leading and lagging strand synthesis. Journal of Biological Chemistry, 266(3):1961-8.

Uchiyama, Y., Takeuchi, R., Kodera, H., and Sakaguchi, K. (2005). Distribution and roles of X-family DNA polymerases in eukaryotes. Biochimie, 91(2), 165-70.

Ulrich, H. D., and Jentsch, S. (2000). Two RING finger proteins mediate cooperation between ubiquitin-conjugating enzymes in DNA repair. EMBO Journal, 19(13), 3388-97.

Ulrich, H. D. (2011). Timing and spacing of ubiquitin-dependent DNA damage bypass. FEBS Letters, 585(18), 2861–2867.

Vaithiyalingam, S., Warren, E. M., Eichman, B. F., and Chazin, W. J. (2010). Insights into eukaryotic DNA priming from the structure and functional interactions of the 4Fe-4S cluster domain of human DNA primase. Proceedings of the National Academy of Sciences, 107(31), 13684-13689.

Van, C., Yan, S., Michael, W. M., Waga, S., and Cimprich, K. A. (2010). Continued primer synthesis at stalled replication forks contributes to checkpoint activation. The Journal of Cell Biology, 189(2), 233–246.

Van der Ploeg, L. H., Schwartz, D. C., Cantor, C. R., and Borst, P. (1984a). Antigenic variation in *Trypanosoma brucei* analyzed by electrophoretic separation of chromosome-sized DNA molecules. Cell, 37(1):77-84.

Van der Ploeg, L. H., Cornelissen, A. W., Barry, J. D., and Borst, P. (1984b). Chromosomes of kinetoplastida. The EMBO Journal, 3(13):3109-15.

Vassin, V. M., Wold, M. S., and Borowiec, J. A. Replication protein A (RPA) phosphorylation prevents RPA association with replication centers. Molecular and Cellular Biology, 24(5), 1930-1943.

Vidal, A. E., Kannouche, P., Podust, V. N., Yang, W., Lehmann, A. R., and Woodgate, R. (2004). Proliferating cell nuclear antigen-dependent coordination of the biological functions of human DNA polymerase ι. The Journal of Biological Chemistry, 279(46), 48360-8.

Vidal, A. E., and Woodgate, R. (2009). Insights into the cellular role of enigmatic DNA polymerase ι. DNA Repair, 8(3), 420-3.

Vincze, T., Posfai, J., and Roberts, R. J. (2003). NEBcutter: A program to cleave DNA with restriction enzymes. Nucleic Acids Research, 31(13), 3688-3691.

Von Hippel, P. H., Fairfield, F. R., and Dolejsi, M. K. (1994). On the processivity of polymerases. Annals of the New York Academy of Sciences, 726, 118–131.

Voss, S., and Skerra, A. (1997) Mutagenesis of a flexible loop in streptavidin leads to higher affinity for the Strep-tag II peptide and improved performance in recombinant protein purification. Protein Engineering Design and Selection, 10(8), 975-982.

Waga, S., and Stillman, B. (1994). Anatomy of a DNA replication fork revealed by reconstitution of SV40 DNA replication in vitro. Nature, 369(6477), 207-12.

Waga, S., and Stillman, B. (1998). The DNA replication fork in eukaryotic cells. Annual Review of Biochemistry, 67, 721-51.

Wanrooij, S., Fusté, J. M., Farge, G., Shi, Y., Gustafsson, C. M., and Falkenberg, M. (2008). Human mitochondrial RNA polymerase primes lagging-strand DNA synthesis *in vitro*. Proceedings of the National Academy of Sciences, 105(32), 11122–11127.

Washington, M. T., Johnson, R. E., Prakash, L., and Prakash, S. (2002). Human DINB1-encoded DNA polymerase κ is a promiscuous extender of mispaired primer termini. Proceedings of the National Academy of Sciences, 99(4), 1910-4.

Ward, J. F. (1975). Radiation-induced strand breakage in DNA. Basic Life Sciences, 5B, 471-2.

Watanabe, K., Tateishi, S., Kawasuji, M., Tsurimoto, T., Inoue, H., and Yamaizumi, M. (2004). Rad18 guides pol η to replication stalling sites through physical interaction and PCNA monoubiquitination. EMBO Journal, 23(19), 3886-96.

Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., and Barton, G. J. (2009). Jalview Version 2 - a multiple sequence alignment editor and analysis workbench. Bioinformatics, 25(9), 1189–1191.

Waters, L. S., and Walker, G. C. (2006). The critical mutagenic translesion DNA polymerase Rev1 is highly expressed during G2/M phase rather than S phase. Proceedings of the National Academy of Sciences, 103(24), 8971–8976.

Weill, J., C., and Reynaud, C., A. (2008). DNA polymerases in adaptive immunity. Nature Reviews Immunology, 8(4), 302-12.

Weller, G. R. and Doherty, A. J. (2001). A family of DNA repair ligases in bacteria? FEBS Letters. 505, 340–342.

Weller, G. R., Kysela, B., Roy, R., Tonkin, L. M., Scanlan, E., Della, M., Devine, S. K., Day, J. P., Wilkinson, A., d'Adda di Fagagna, F., Devine, K. M., Bowater, R. P., Jeggo, P. A., Jackson, S. P., and Doherty, A. J. (2002). Identification of a DNA nonhomologous end-joining complex in bacteria. Science, 297(5587):1686-9.

Wickner, W., Brutlag, D., Schekman, R., and Kornberg, A. (1972). RNA synthesis initiates in vitro conversion of M13 DNA to its replicative form. Proceedings of the National Academy of Sciences, 69(4), 965-9.

Wist, E., and Prydz, H. (1979). The effect of aphidicolin on DNA synthesis in isolated HeLa cell nuclei. Nucleic Acids Research, 6(4), 1583-1590.

Wittschieben, J., Shivji, M. K., Lalani, E., Jacobs, M. A., Marini, F., Gearhart, P. J., Rosewell, I., Stamp, G., and Wood, R. D. (2000). Disruption of the developmentally regulated Rev3L gene causes embryonic lethality. Current Biology, 10(19), 1217-20.

Wobbe, C. R., Weissbach, L., Borowiec, J. A., Dean, F. B., Murakami, Y., Bullock, P., and Hurwitz, J. (1987). Replication of simian virus 40 origin-containing DNA in vitro with purified proteins. Proceedings of the National Academy of Sciences, 84(7), 1834-8.

Wold, M. S., and Kelly, T. (1988). Purification and characterization of replication protein A, a cellular protein required for in vitro replication of simian virus 40 DNA. Proceedings of the National Academy of Sciences, 85(8), 2523-7.

Woodward, R., and Gull, K. (1990). Timing of nuclear and kinetoplast DNA replication and early morphological events in the cell cycle of *Trypanosoma brucei*. Journal of Cell Science, 95(1), 49–57.

World Health Organisation. (2008). The global burden of disease: 2004 update. (Switzerland, WHO Press).

Wong, T. W., and Clayton, D. A. (1985). Isolation and characterization of a DNA primase from human mitochondria. The Journal of Biological Chemistry, 260(21), 11530–11535.

Wong, T. W., and Clayton, D. A. (1986). DNA primase of human mitochondria is associated with structural RNA that is essential for enzymatic activity, Cell, 45(6), 817–825.

Wright, G. E., Hubscher, U., Khan, N. N., Focher, F., and Verri, A. 1994. Inhibitor analysis of calf thymus DNA polymerases alpha, delta and epsilon. FEBS Letters, 341(1), 128-130.

Xu, B., and Clayton, D. A. (1996). RNA-DNA hybrid formation at the human mitochondrial heavy-strand origin ceases at replication start sites: An implication for RNA-DNA hybrids serving as primers. The EMBO Journal, 15, 3135–3143.

Yang, X. H., Shiotani, B., Classon, M. and Zou, L. (2008). Chk1 and Claspin potentiate PCNA ubiquitination. Genes and Development, 22, 1147–1152.

Yeeles, J. T. P., and Marians, K. J. (2011). The *Escherichia coli* Replisome Is Inherently DNA Damage Tolerant. Science, 334(6053), 235–238.

Yoon, J. H., Prakash, L., and Prakash, S. (2010a). Error-free replicative bypass of (6-4) photoproducts by DNA polymerase ζ in mouse and human cells. Genes and Development, 24(2), 123-8.

Yoon, J. H., Bhatia, G., Prakash, S., and Prakash, L. (2010b). Error-free replicative bypass of thymine glycol by the combined action of DNA polymerases κ and ζ in human cells. Proceedings of the National Academy of Sciences, 107(32), 14116-21.

Young, C. W., and Hodas, S. (1964). Hydroxyurea: inhibitory effect on DNA metabolism. Science, 146(3648), 1172-1174.

Zeng, X., Winter, D. B., Kasmer, C., Kraemer, K. H., Lehmann, A. R., and Gearhart, P. J. (2001). DNA polymerase η is an A-T mutator in somatic hypermutation of immunoglobulin variable genes. Nature Immunology, 2(6), 537-41.

Zhang, H., Chatterjee, A., and Singh, K. K. (2006). *Saccharomyces cerevisiae* polymerase ζ functions in mitochondria. Genetics, 172(4), 2683-2688.

Zhang, Y., Yuan, F., Wu, X., and Wang, Z. (2000a). Preferential incorporation of G opposite template T by the low-fidelity human DNA polymerase ι. Molecular and Cellular Biology, 20(19), 7099–7108.

Zhang, Y., Yuan, F., Wu, X., Wang, M., Rechkoblit, O., Taylor, J. S., Geacintov, N. E., and Wang, Z. (2000b). Error-free and error-prone lesion bypass by human DNA polymerase κ *in vitro*. Nucleic Acids Research, 28(21), 4138-46.

Zhang, Y., Yuan, F., Wu, X., Rechkoblit, O., Taylor, J. S., Geacintov, N. E., and Wang, Z. (2000c). Error-prone lesion bypass by human DNA polymerase η. Nucleic Acids Research, 28 4717– 4724.

Zhang, Y., Wu, X., Rechkoblit, O., Geacintov, N. E., Taylor, J. S., and Wang, Z. (2002). Response of human REV1 to different DNA damage: preferential dCMP insertion opposite the lesion. Nucleic Acids Research, 30(7), 1630-8.

Zheng, L., Baumann, U., and Reymond, J. L. (2004). An efficient one-step site-directed and site-saturation mutagenesis protocol. Nucleic Acids Research, 32(14), e115.

Zheng, L., Zhou, M., Guo, Z., Lu, H., Qian, L., Dai, H., Qiu, J., Yakubovskaya, E., Bogenhagen, D. F., Demple, B., and Shen, B. (2008). Human DNA2 is a mitochondrial nuclease/helicase for efficient processing of DNA replication and repair intermediates. Molecular Cell, 32(3), 325-336.

Zhong, X., Garg, P., Stith, C. M., Nick, McElhinny, S. A., Kissling, G. E., Burgers, P. M., and Kunkel, T. A. (2006). The fidelity of DNA synthesis by yeast DNA polymerase ζ alone and with accessory proteins. Nucleic Acids Research, 34(17), 4731-42.

Zhu, H., Nandakumar, J., Aniukwu, J., Wang, L. K., Glickman, M. S., Lima, C. D., and Shuman, S. (2006). Atomic structure and nonhomologous end-joining function of the polymerase component of bacterial DNA ligase D. Proceedings of the National Academy of Sciences, 103(6), 1711-6.

Ziv, O., Geacintov, N., Nakajima, S., Yasui, A., and Livneh, Z. (2009). DNA polymerase ζ cooperates with polymerases κ and ι in translesion DNA synthesis across pyrimidine photodimers in cells from XPV patients. Proceedings of the National Academy of Sciences, 106(28), 11552–11557.

Ziv, O., Diamant, N., Shachar, S., Hendel, A., and Livneh, Z. (2012). Quantitative measurement of translesion DNA synthesis in mammalian cells. Methods in Molecular Biology, 920, 529-542.

Zlatanou, A., Despras, E., Braz-Petta, T., Boubakour-Azzouz, I., Pouvelle, C., Stewart, G. S., Nakajima, S., Yasui, A., Ishchenko, A. A., and Kannouche, P. L. (2011). The hMsh2-hMsh6 complex acts in concert with monoubiquitinated PCNA and Pol η in response to oxidative DNA damage in human cells. Molecular Cell, 43(4), 649-62.

Appendix

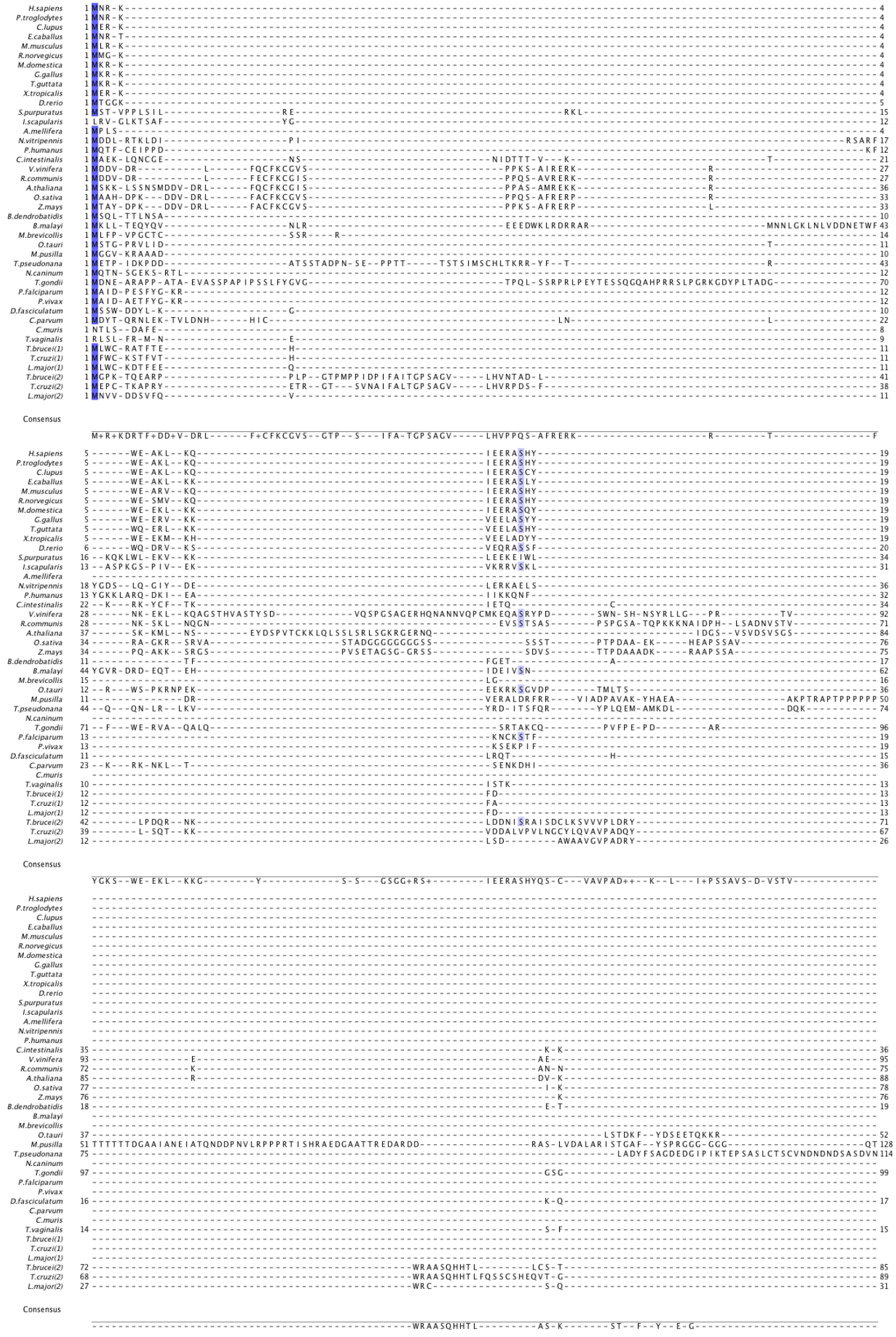
Multiple sequence alignment of PrimPol homologues from a broad range of eukaryotic species, including trypanosomatid PrimPol1 and PrimPol2 families.

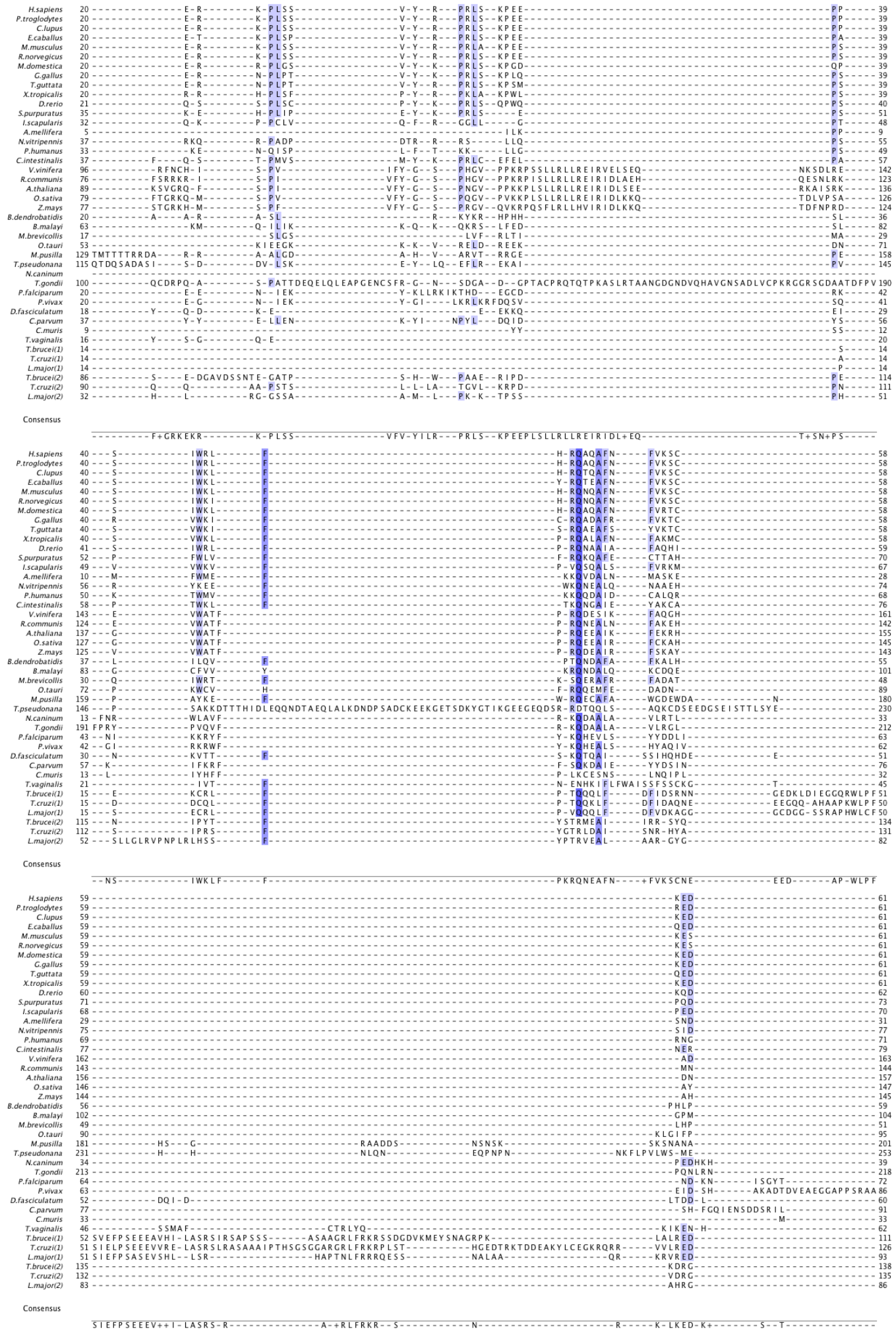
The accession numbers of homologues used are as follows:
Homo sapien, NP_689896.1; *Pan troglodytes*, XP_001162592.1; *Macaca mulatta*, XP_001083404.2; *Equus caballus*, XP_001491218; *Canis lupis familiaris*, XP_532846.3; *Monodelphis domestica*, XP_001369409; *Rattus norvegicus*, XP_341434.3; *Mus musculus*, NP_001001184.1; *Taeniopygia guttata*, XP_002190203; *Danio rerio*, NP_001032455.1; *Strongylocentrotus purpuratus*, XP_785583; *Ciona intestinalis*, XP_002123260; *Monosigna brevicollis*, XP_001749128; *Ricinus communis*, XP_002513641; Vitis vinifera, XP_002274252; *Arabidopsis thaliana*, AED96263; *Zea Mays*, NP_001141209; *Oryza sativa japonica* BAG93719; *Batrachochytrium dendrobatidis*, EGF82076; *Micromonas pusilla*, XP_003063911; *Thalassiosira pseudonana*, XP_002287384; *Brugia malayi*, XP_001898035; *Ostreococcus tauri*, XP_003081466; *Cryptosporidium parvum*, XP_626834; *Ixodes scapularis*, XP_002402586; *Nasonia vitripennis*, XP_001603706; *Pediculus humanus* XP_002428759; *Apis mellifera* XP_001121815; *Dictyostelium discoideum*, XP_646915; *Trichomonas vaginalis,*XP_001305445

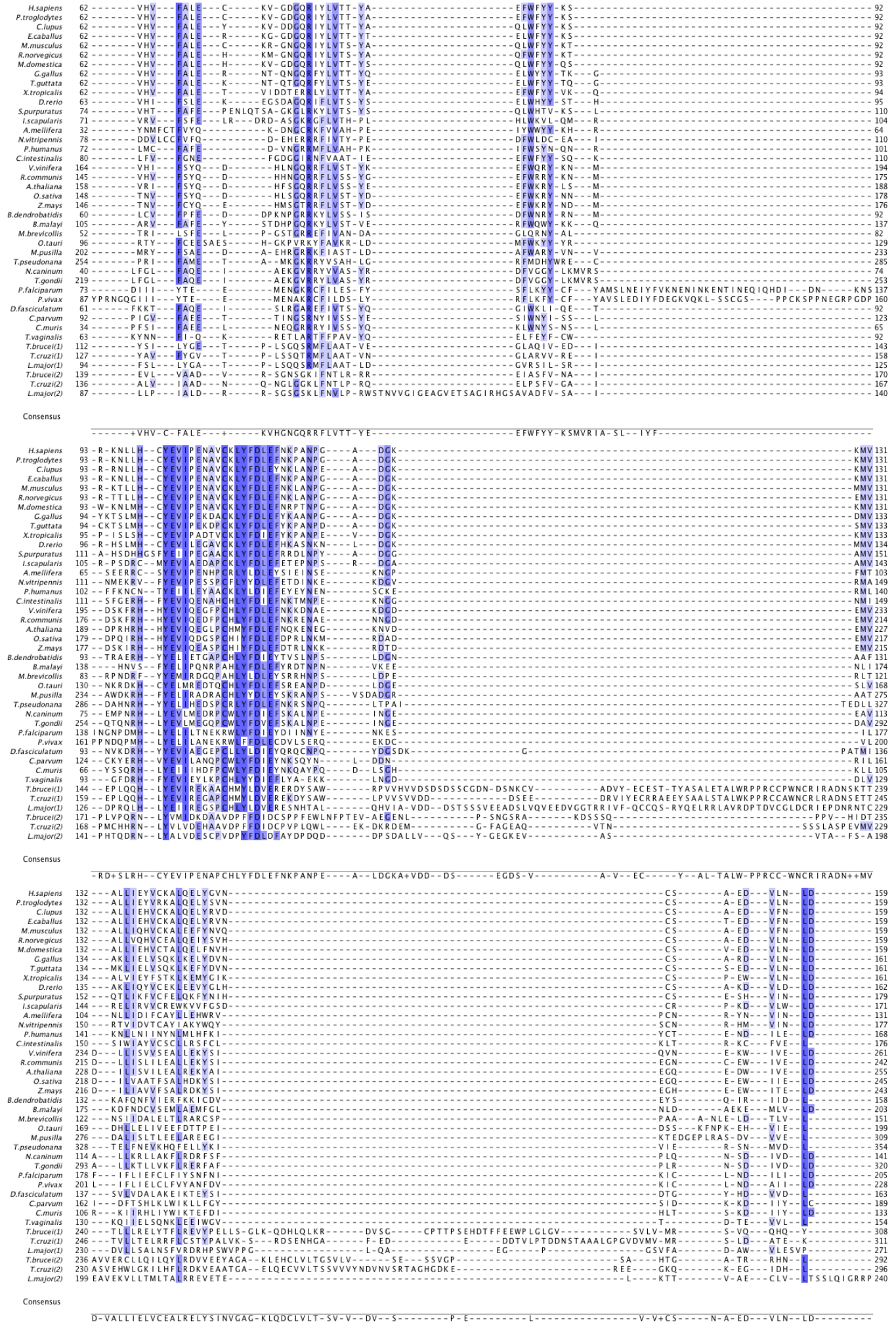
PrimPol1, indicated by (1): *Trypanosoma cruzi*, XP_812815; *Trypanosoma brucei*, XP_845077; *Leishmania major* XP_001685728

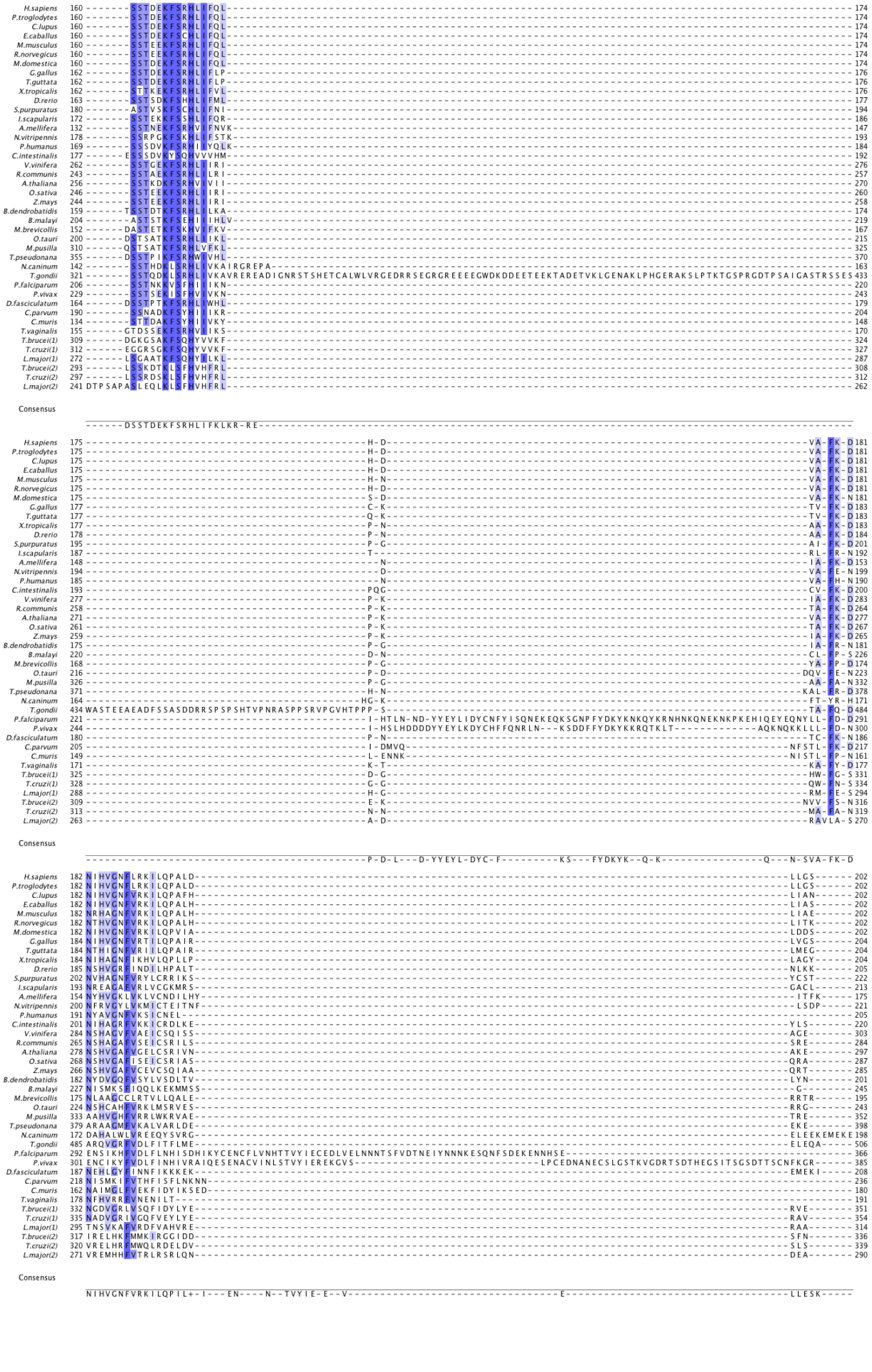
PrimPol2, indicated by (2): *Leishmania major,* XP_001686100; *Trypanosoma brucei,* XP_822505.1; *Trypanosoma cruzi*, XP_819101

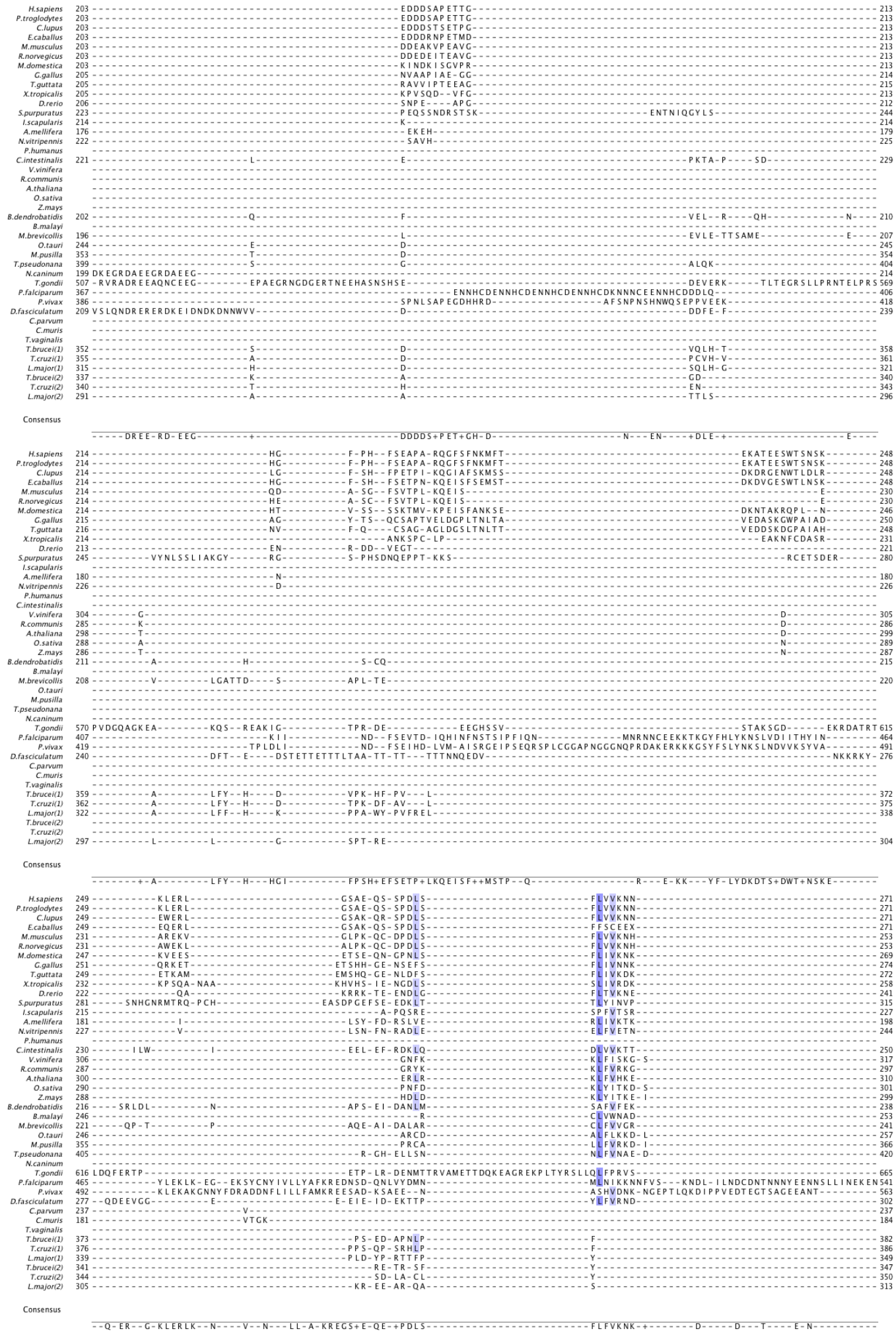
Alignment was generated using ClustalW2 (Larkin *et al*., 2007), TCoffee (Notredame *et al*., 2007), MUSCLE (Edgar, 2004), and edited in JALview (Waterhouse *et al*., 2009). Blue shading indicates ≥40 % sequence identity.

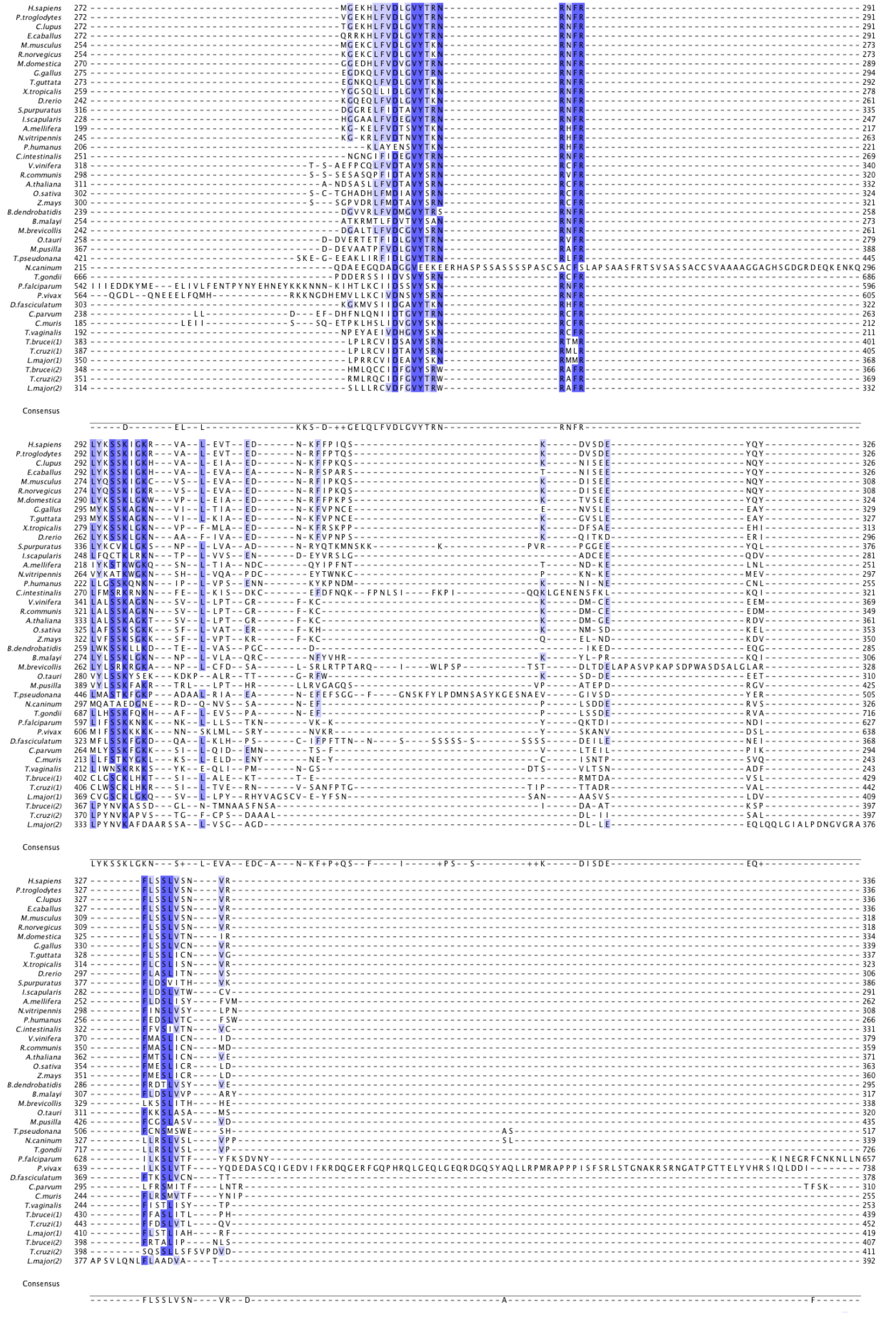

H.sapiens
P.troglodytes
C.lupus
E.caballus
M.musculus
R.norvegicus
M.domestica
G.gallus
T.guttata
X.tropicalis
D.rerio
S.purpuratus
I.scapularis
A.mellifera
N.vitripennis
P.humanus
C.intestinalis
V.vinifera
R.communis
A.thaliana
O.sativa
Z.mays
B.dendrobatidis
B.malayi
M.brevicollis
O.tauri
M.pusilla
T.pseudonana
N.caninum
T.gondii
P.falciparum
P.vivax
D.fasciculatum
C.parvum
C.muris
T.vaginalis
T.brucei(1)
T.cruzi(1)
L.major(1)
T.brucei(2)
T.cruzi(2)
L.major(2)
Consensus

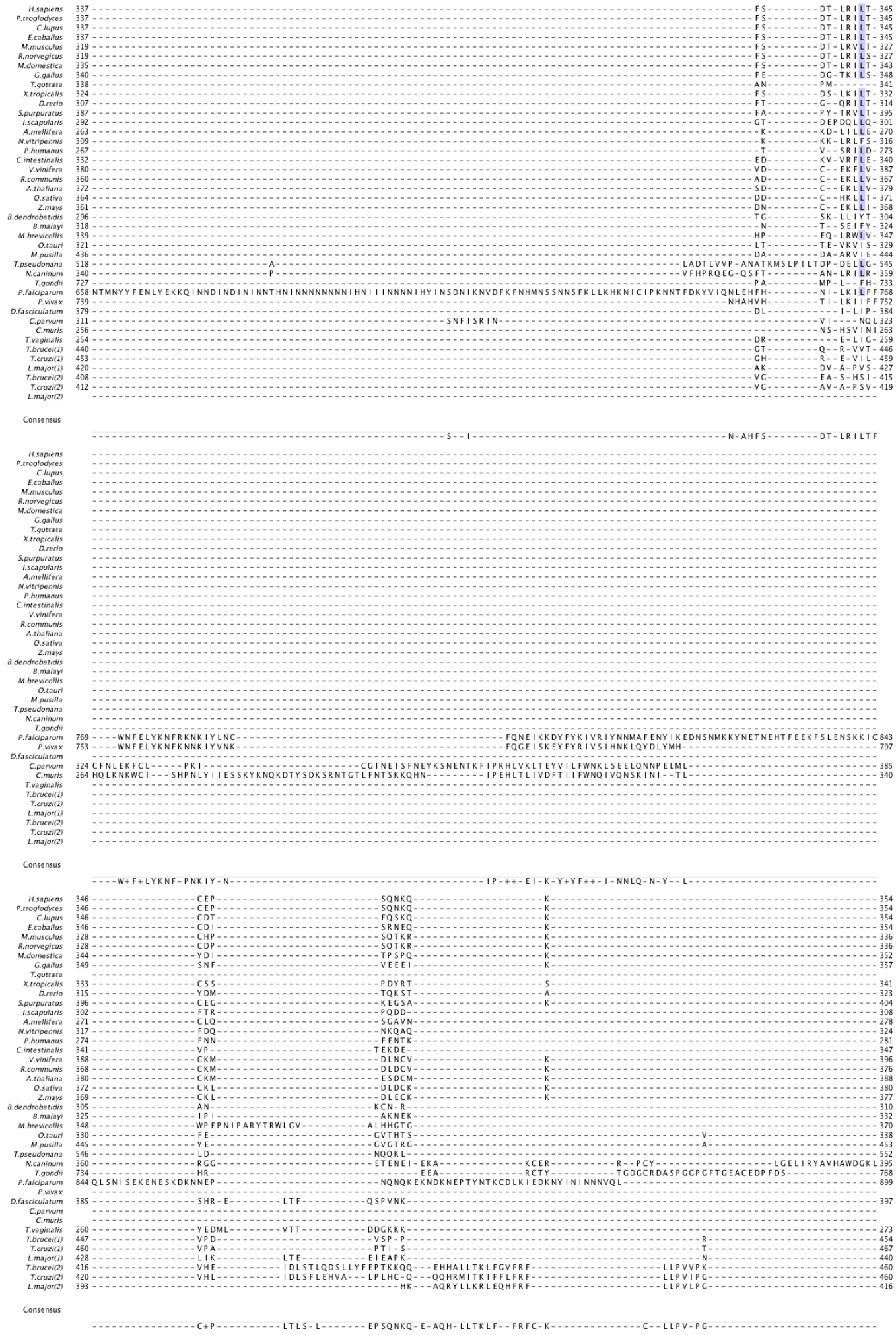


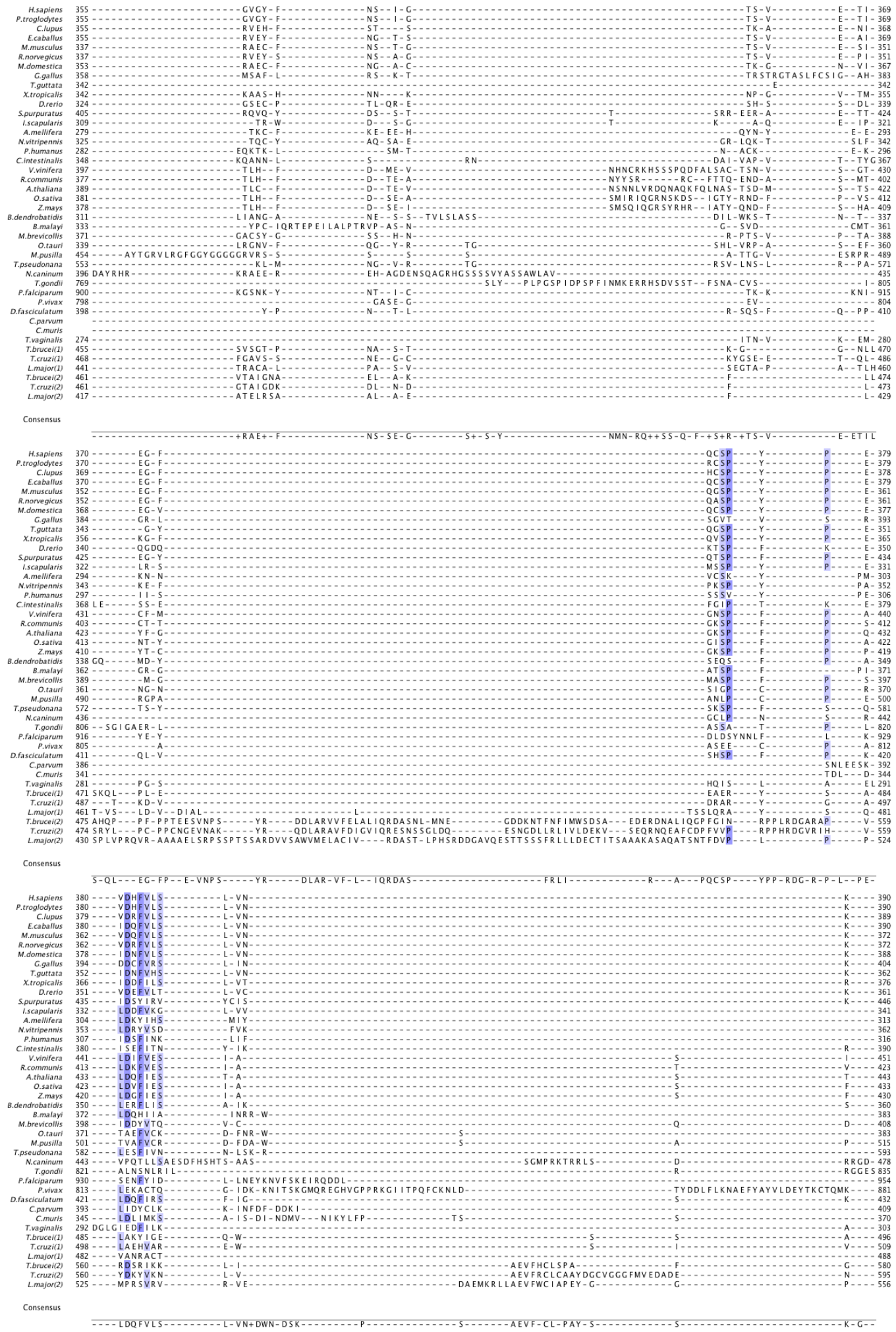

H.sapiens
P.troglodytes
C.lupus
E.caballus
M.musculus
R.norvegicus
M.domestica
G.gallus
T.guttata
X.tropicalis
D.rerio
S.purpuratus
I.scapularis
A.mellifera
N.vitripennis
P.humanus
C.intestinalis
V.vinifera
R.communis
A.thaliana
O.sativa
Z.mays
B.dendrobatidis
B.malayi
M.brevicollis
O.tauri
M.pusilla
T.pseudonana
N.caninum
T.gondii
P.falciparum
P.vivax
D.fasciculatum
C.parvum
C.muris
T.vaginalis
T.brucei(1)
T.cruzi(1)
L.major(1)
T.brucei(2)
T.cruzi(2)
L.major(2)
Consensus

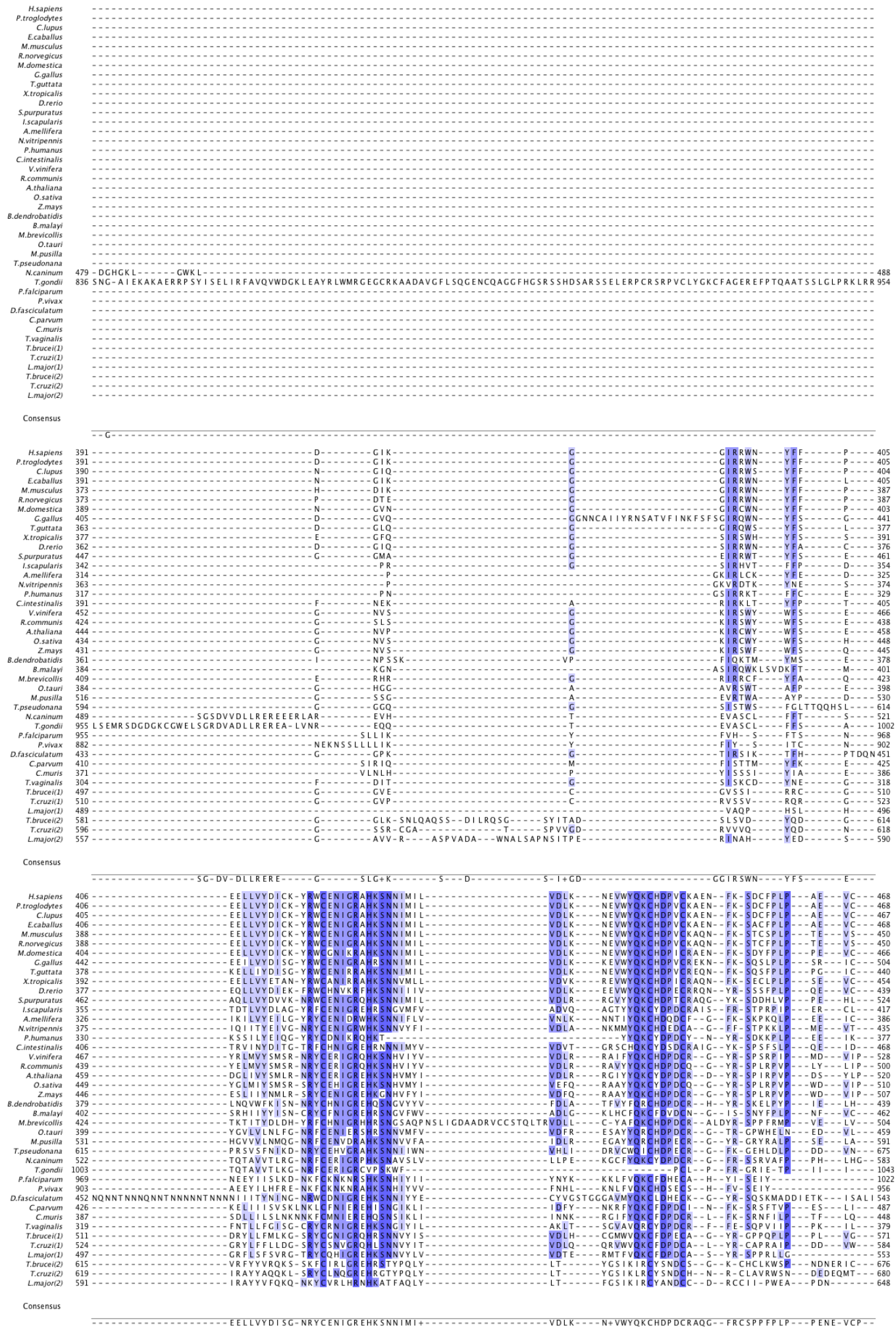

H.sapiens
P.troglodytes
C.lupus
E.caballus
M.musculus
R.norvegicus
M.domestica
G.gallus
T.guttata
X.tropicalis
D.rerio
S.purpuratus
I.scapularis
A.mellifera
N.vitripennis
P.humanus
C.intestinalis
V.vinifera
R.communis
A.thaliana
O.sativa
Z.mays
B.dendrobatidis
B.malayi
M.brevicollis
O.tauri
M.pusilla
T.pseudonana
N.caninum
T.gondii
P.falciparum
P.vivax
D.fasciculatum
C.parvum
C.muris
T.vaginalis
T.brucei(1)
T.cruzi(1)
L.major(1)
T.brucei(2)
T.cruzi(2)
L.major(2)
Consensus

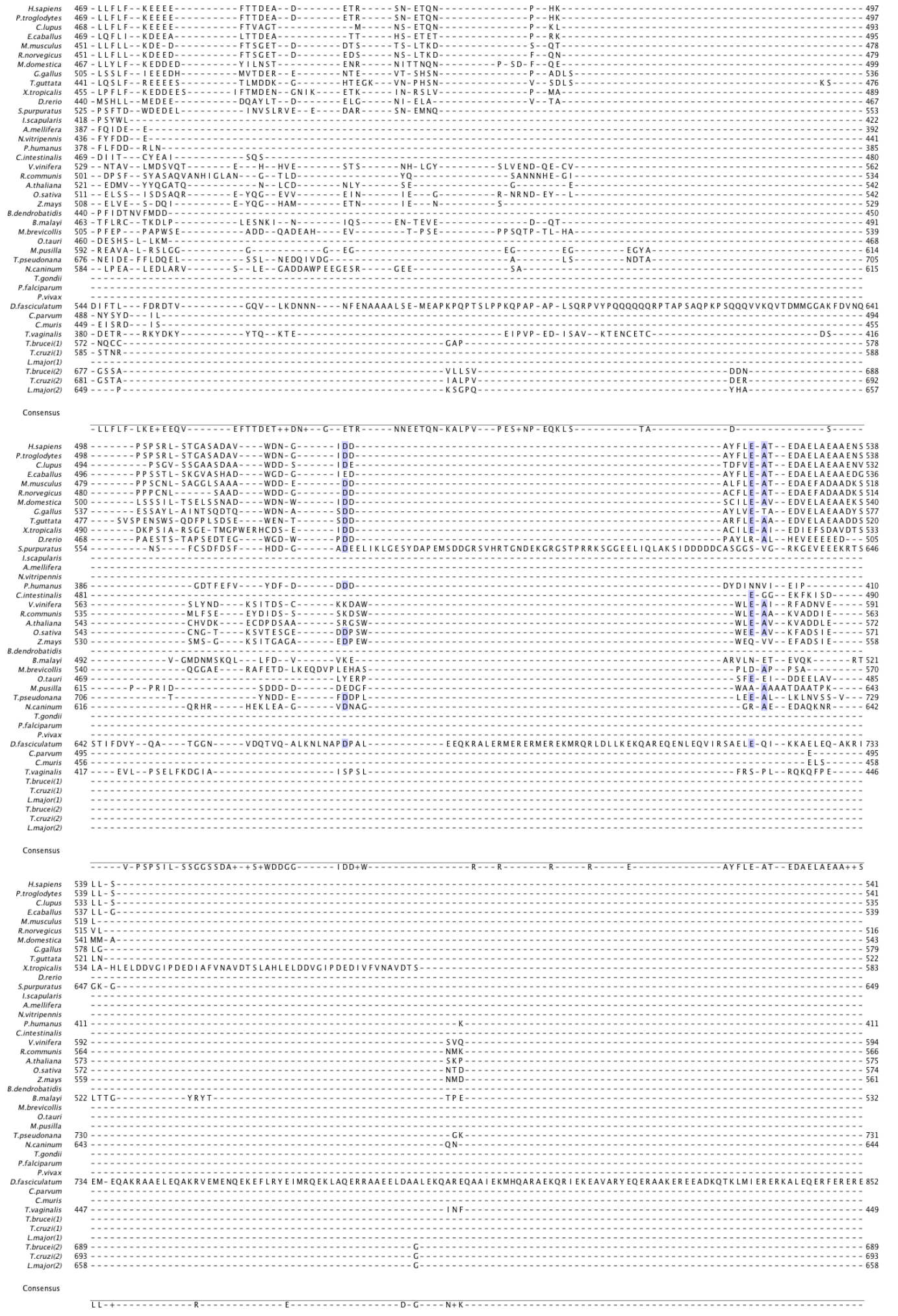


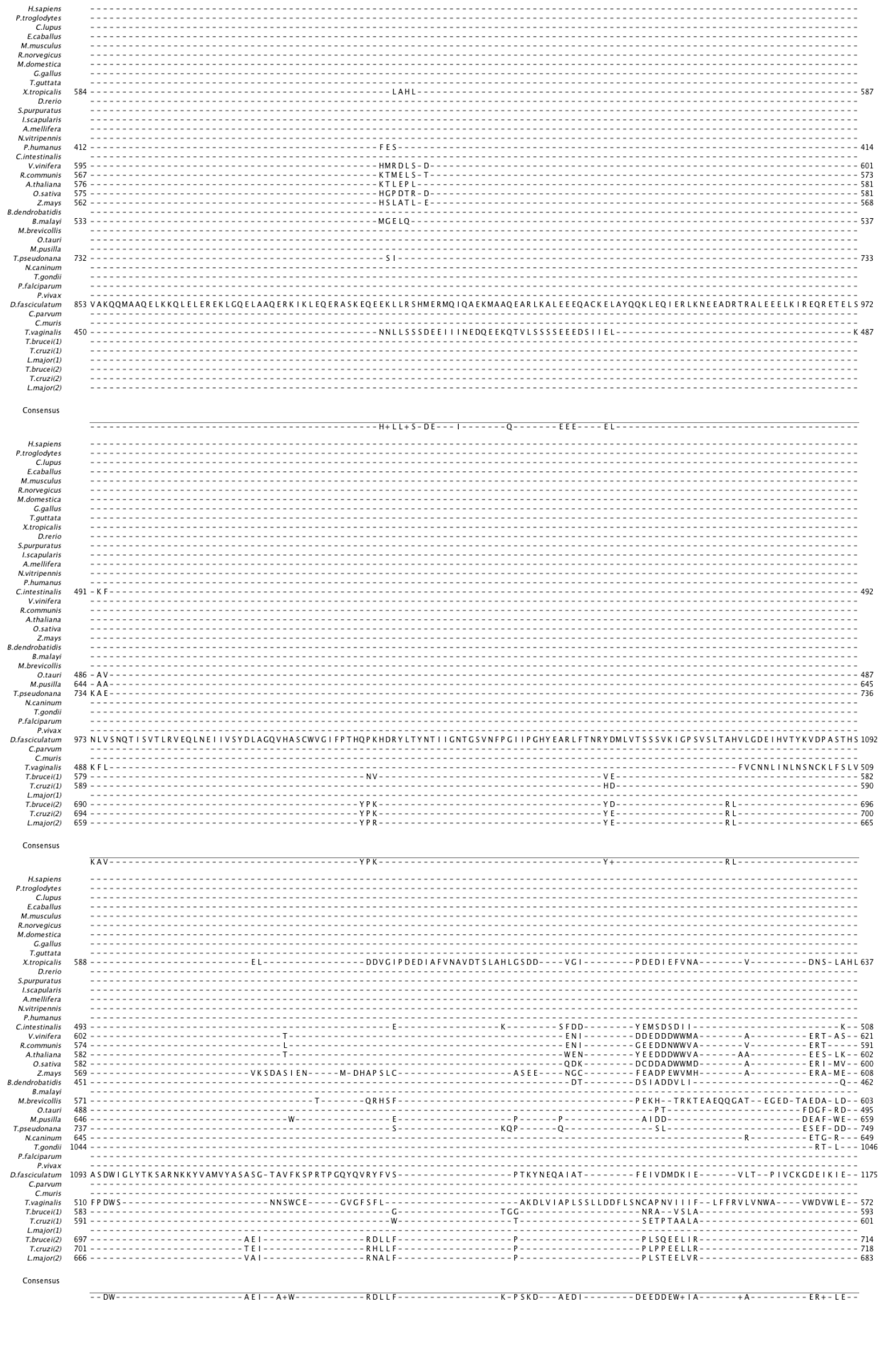


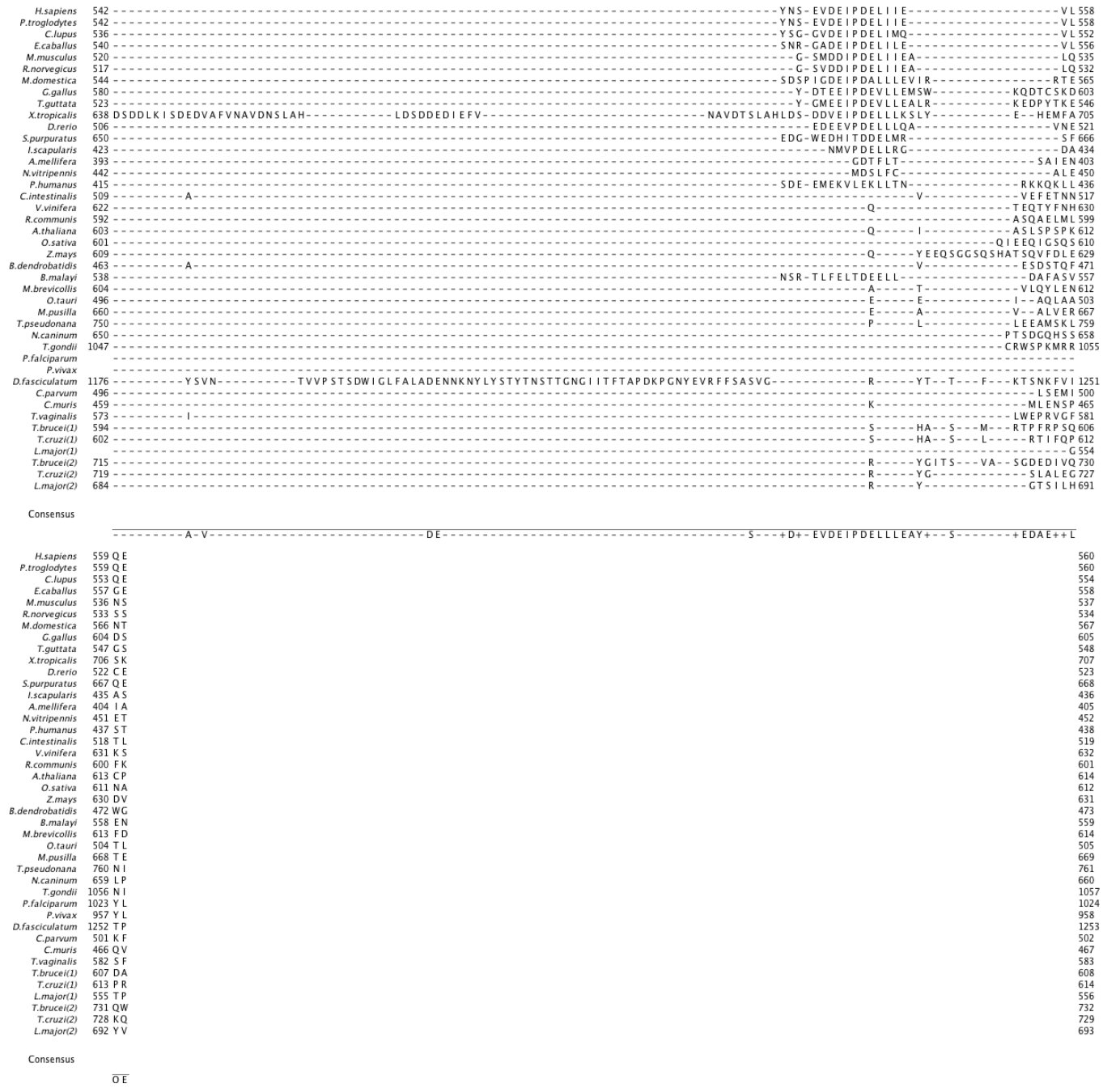





H.sapiens
P.troglodytes
C.lupus
E.caballus
M.musculus
R.norvegicus
M.domestica
G.gallus
T.guttata
X.tropicalis
D.rerio
S.purpuratus
I.scapularis
A.mellifera
N.vitripennis
P.humanus
C.intestinalis
V.vinifera
R.communis
A.thaliana
O.sativa
Z.mays
B.dendrobatidis
B.malayi
M.brevicollis
O.tauri
M.pusilla
T.pseudonana
N.caninum
T.gondii
P.falciparum
P.vivax
D.fasciculatum
C.parvum
C.muris
T.vaginalis
T.brucei(1)
T.cruzi(1)
L.major(1)
T.brucei(2)
T.cruzi(2)
L.major(2)
Consensus




H.sapiens
P.troglodytes
C.lupus
E.caballus
M.musculus
R.norvegicus
M.domestica
G.gallus
T.guttata
X.tropicalis
D.rerio
S.purpuratus
I.scapularis
A.mellifera
N.vitripennis
P.humanus
C.intestinalis
V.vinifera
R.communis
A.thaliana
O.sativa
Z.mays
B.dendrobatidis
B.malayi
M.brevicollis
O.tauri
M.pusilla
T.pseudonana
N.caninum
T.gondii
P.falciparum
P.vivax
D.fasciculatum
C.parvum
C.muris
T.vaginalis
T.brucei(1)
T.cruzi(1)
L.major(1)
T.brucei(2)
T.cruzi(2)
L.major(2)
Consensus


H.sapiens 542 YNS-EVDEIPDELIIE 558
P.troglodytes 542 YNS-EVDEIPDELIIE 558
C.lupus 536 YSG-GVDEIPDELIMQ 552
E.caballus 540 SNR-GADEIPDELILE 556
M.musculus 520 G-SMDDIPDELIIEA 535
R.norvegicus 517 G-SVDDIPDELIIEA 532
M.domestica 544 SDSPIGDEIPDALLLEVIR 565
G.gallus 580 Y-DTEEIPDEVLLEMSW 603
T.guttata 523 Y-GMEEIPDEVLLEALR 546
X.tropicalis 638 DSDDLKISDEDVAFVNAVDNSLAH LDSDDEDIEFV NAVDTSLAHLDS-DDVEIPDELLLKSLY E--HEMFA 705
D.rerio 506 EDEEVPDELLLQA 521
S.purpuratus 650 EDG-WEDHITDDELMR 666
I.scapularis 423 NMVPDELLRG 434
A.mellifera 393 GDTFLT 403
N.vitripennis 442 MDSLFC 458
P.humanus 415 SDE-EMEKVLEKLLTN RKKQKLL 436
C.intestinalis 509 A V 517
V.vinifera 622 Q TEQTYFNH 630
R.communis 592 ASQAELML 599
A.thaliana 603 Q I ASLSPSPK 612
O.sativa 601 QIEEQIGSQS 610
Z.mays 609 Q YEEQSGGSQSHATSQVFDLE 629
B.dendrobatidis 463 A V ESDSTQF 471
B.malayi 538 NSR-TLFELTDEELL DAFASV 557
M.brevicollis 604 A T VLQYLEN 612
O.tauri 496 E E I--AQLAA 503
M.pusilla 660 E A V--ALVER 667
T.pseudonana 750 P L LEEAMSKL 759
N.caninum 650 PTSDGQHSS 658
T.gondii 1047 CRWSPKMRR 1055
P.falciparum
P.vivax
D.fasciculatum 1176 YSVN TVVPSTSDWIGLFALADENNKNYLYSTYTNSTTGNGIITFTAPDKPGNYEVRFFSASVG R YT--T---F---KTSNKFVI 1251
C.parvum 496 LSEMI 500
C.muris 459 K MLENSP 465
T.vaginalis 573 I LWEPRVGF 581
T.brucei(1) 594 S HA--S---M---RTPFRPSQ 606
T.cruzi(1) 602 S HA--S---L-----RTIFQP 612
L.major(1) G 554
T.brucei(2) 715 R YGITS---VA--SGDEDIVQ 730
T.cruzi(2) 719 R YG SLALEG 727
L.major(2) 684 R Y GTSILH 691
Consensus
A-V DE S---+D+-EVDEIPDELLLEAY+--S EDAE++L
H.sapiens 559 QE 560
P.troglodytes 559 QE 560
C.lupus 553 QE 554
E.caballus 557 GE 558
M.musculus 536 NS 537
R.norvegicus 533 SS 534
M.domestica 566 NT 567
G.gallus 604 DS 605
T.guttata 547 GS 548
X.tropicalis 706 SK 707
D.rerio 522 CE 523
S.purpuratus 667 QE 668
I.scapularis 435 AS 436
A.mellifera 404 IA 405
N.vitripennis 451 ET 452
P.humanus 437 ST 438
C.intestinalis 518 TL 519
V.vinifera 631 KS 632
R.communis 600 FK 601
A.thaliana 613 CP 614
O.sativa 611 NA 612
Z.mays 630 DV 631
B.dendrobatidis 472 WG 473
B.malayi 558 EN 559
M.brevicollis 613 FD 614
O.tauri 504 TL 505
M.pusilla 668 TE 669
T.pseudonana 760 NI 761
N.caninum 659 LP 660
T.gondii 1056 NI 1057
P.falciparum 1023 YL 1024
P.vivax 957 YL 958
D.fasciculatum 1252 TP 1253
C.parvum 501 KF 502
C.muris 466 QV 467
T.vaginalis 582 SF 583
T.brucei(1) 607 DA 608
T.cruzi(1) 613 PR 614
L.major(1) 555 TP 556
T.brucei(2) 731 QW 732
T.cruzi(2) 728 KQ 729
L.major(2) 692 YV 693
Consensus
O E